

Influence of the Chromophore-Protein-Interaction on the
chromophore isomerisation in Photoactive Yellow Protein

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung des Grades
einer Doktorin der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von Helga Maria Maike Clemens
aus Mönchengladbach

1. Referent: Professor Dr. Peter Jomo Walla
2. Referent: Dr. Gerrit Groenhof
eingereicht am: 02.09.2013
mündliche Prüfung (Disputation) am: 20.05.2014

Druckjahr 2014

Veröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Posterbeiträge

Clemens, M. & Groenhof, G.: Excited state dynamics simulations of photoactive yellow protein: the influence of QM/MM partitioning and choice of active space orbitals, Poster 014, Computational Simulations and Theory of Macromolecules (2011).

Clemens, M. & Inhester, L. & Groenhof, G.: Short Hydrogen Bonds Connecting the Chromophore and Glu46 in Photoactive Yellow Protein, Poster 131, Computational Simulations and Theory of Macromolecules (2012).

Clemens, M. & Davari Dolatabadi, M. & Groenhof, G.: Hydrogen bonds around the chromophore of photoactive yellow protein, Poster 251, Annual Meeting of the German Biophysical Society (2012).

Contents

1	Introduction	1
1.1	Photoactive Yellow Protein	1
1.1.1	Chromophore Deactivation	5
1.2	Hydrogen Bonds	8
1.3	Motivation	13
2	Methods & Theory	15
2.1	Molecular Dynamics	15
2.2	Molecular Mechanics	18
2.3	Hybrid Quantum Mechanical/Molecular Mechanics	20
2.4	Density Functional Theory	23
2.5	Complete Active Space Self Consistent Field (CASSCF)	27
2.6	Conical Intersection	37
2.6.1	Surface Hopping	40
3	Energy Potential Along the Chromophore - Glu46 Hydrogen Bond	41
3.1	Influences on the Hydrogen Bond Energy Potential	43
3.2	Location of Hydrogen Atom	53
3.3	Minimised Structures	55
3.4	Discussion	57
4	Influence of Arg52 Protonation State on PYP Dynamics and Chromophore Isomerisation	61
4.1	Dynamics	64
4.1.1	QMMM Dynamics	76
4.2	Discussion Ground State MD Simulations	81
4.3	Deactivation Events	85
4.4	Discussion of Excited State Deactivations	99

5	pCK⁻ Deactivation in Water and Decanol	107
5.1	pCK ⁻ in Water	108
5.2	pCK ⁻ in Decanol	111
5.3	Discussion	113
6	Conclusion & Outlook	117
7	Appendix	123
	List of Figures	141
	List of Tables	148
	Bibliography	152

Chapter 1

Introduction

Reacting to changes of the environment is essential to all organisms. One important environmental factor is light. It can be beneficial, as in the case of photosynthesis, or harmful when it comes to wave length and light intensities that would destroy organisms.

To enable an organism to react to light in an appropriate manner, proteins are needed that are sensitive to wave length or light intensities and to transduce this information such that the organism is able to react. This task is accomplished by proteins to which a chromophore is bound. Light absorption of the chromophore induces conformational changes of the light sensing protein or chemical processes that are detected by other proteins. This way the absorbed light energy is converted to structural or chemical changes.

One light sensing protein is Photoactive Yellow Protein, the subject of investigation in this thesis.

1.1 Photoactive Yellow Protein

Photoactive Yellow Protein (PYP) is a small globular protein, which was discovered in *Halorhodospira halophila* in 1985¹. It is the first element of a protein cascade that leads to negative phototaxis, the bacterial movement away from a harmful blue light source². PYP, which is depicted in Figure 1.1 consists of 125 amino acids that form a central six stranded β -sheet, which is surrounded by five short α -helices. The chromophore *p*-coumaric acid (pCA) is bound via a thio-ester linkage to Cys69³.

PYP is the structure prototype for the Per-Arnt-Sim domain (PAS domain) super family⁴⁻⁶, which comprises over 1300 proteins⁷. The PAS domain super family is an important group of proteins because they sense changes in the

environment, such as light, concentrations of ligands or redox potentials, and convey this information to other proteins⁸. The PAS domain consists of a five-stranded β -sheet whose strands are interconnected by α -helices. These α -helices form a pocket with the β -sheet facilitating ligand binding. In case of PYP, this is the chromophore binding pocket^{4,9–11}.

Combining these facts with the relative ease with which PYP can be synthesised and investigated experimentally makes it a model system for PAS domains.

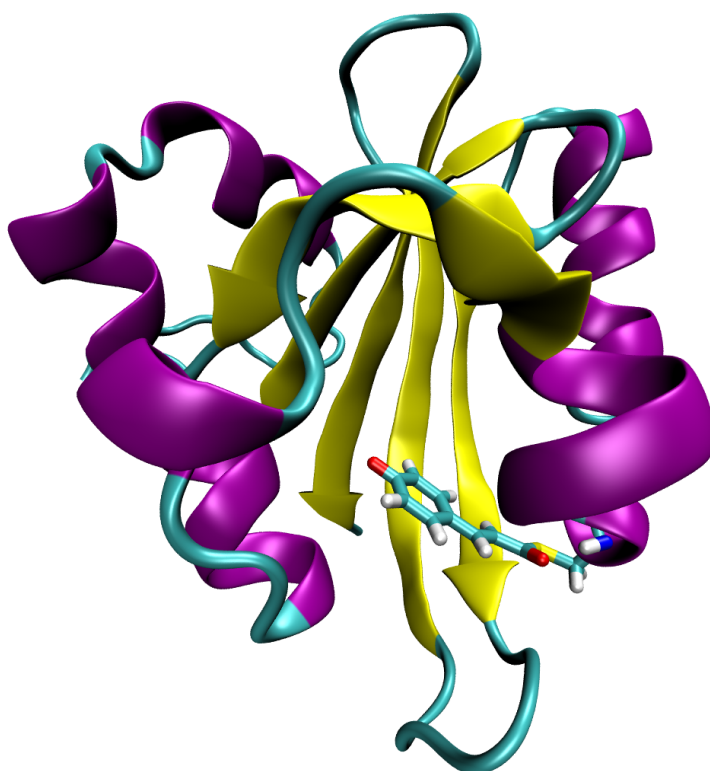


Figure 1.1: Structure of PYP with the position of the chromophore bound to Cys69 via a thio-ester linkage (depicted as explicit atoms). The β -sheet is shown in yellow whereas the connecting α -helices are coloured purple.

The chromophore binding pocket encloses the negatively charged chromophore completely inside the protein and is depicted in Figure 1.2. Beside the covalent thio-ester linkage to Cys69, the chromophore forms hydrogen bonds to the neighbouring amino acids Tyr42, Glu46, and Cys69^{3,12–14}. These hydrogen bonds help to stabilise the negative charge that is delocalised via conjugated double bonds over the phenyl ring and the ethylene chain.

Arg52 separates the chromophore from the bulk water and forms hydrogen bonds to Thr50 and Tyr98. These hydrogen bonded structure prevents bulk water molecules from entering the chromophore binding pocket.

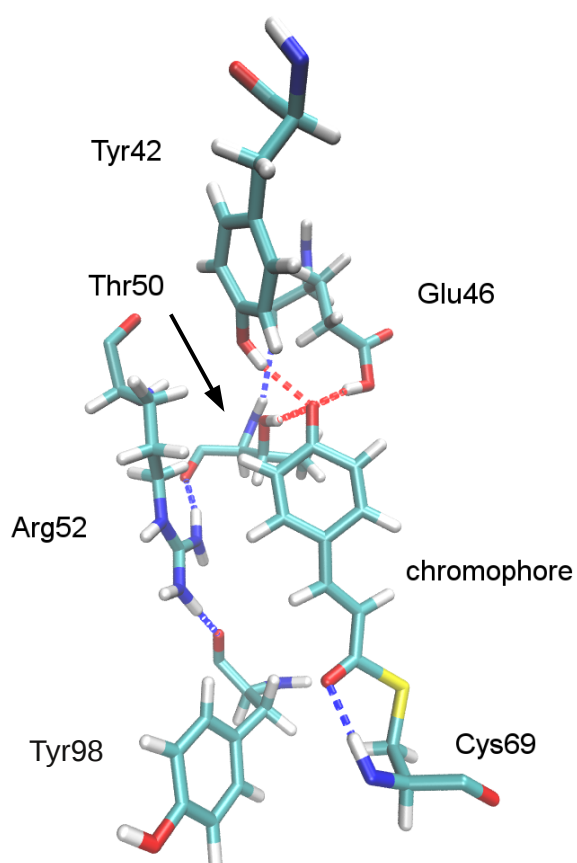


Figure 1.2: The PYP chromophore surrounded by the amino acids that form the binding pocket. The chromophore forms hydrogen bonds to Tyr42, Glu46, Thr50, and Cys69. Arg52 and Thr50 and Tyr98 are connected by a hydrogen bond among themselves.

PYP undergoes a photocycle upon light excitation¹⁵, which is depicted in Figure 1.3. In the ground state the chromophore is trans-configured (pG state). After excitation by blue light (446 nm), it deactivates from the excited

state via isomerisation to a *cis*-configuration (pR state)^{13,14}. After isomerisation, which changes the chromophore's pK_a , it is protonated by the hydrogen of the Glu46 carboxyl group to which it is still hydrogen bonded^{16,17}. At this point of the photocycle, Arg52, which separates the chromophore from the bulk water in ground state, functions as a lid to the chromophore binding pocket. Its sidechain flips open and allows bulk water to enter the binding pocket^{18–21}. These water molecules are believed to stabilise the protonated chromophore and to help during the recovery of the ground state^{13,15}.

The state with protonated chromophore and flipped Arg52 is called the signalling state (pB) of PYP. Large configurational changes aside from the chromophore isomerisation and the Arg52 movement were not detected during the photocycle in diffraction experiments^{18,22–24}. FTIR experiments comparing PYP movements during the photocycle in solution and crystallised state show that only small changes occur in the crystal lattice compared with the protein in solution¹⁶. This result was expected but should still be kept in mind when comparing data obtained from PYP in solution with X-ray structure analysis.

A detailed description of the primary events of the photocycle with emphasis on the deactivation pathway from the excited state chromophore will be presented in the following subsection.

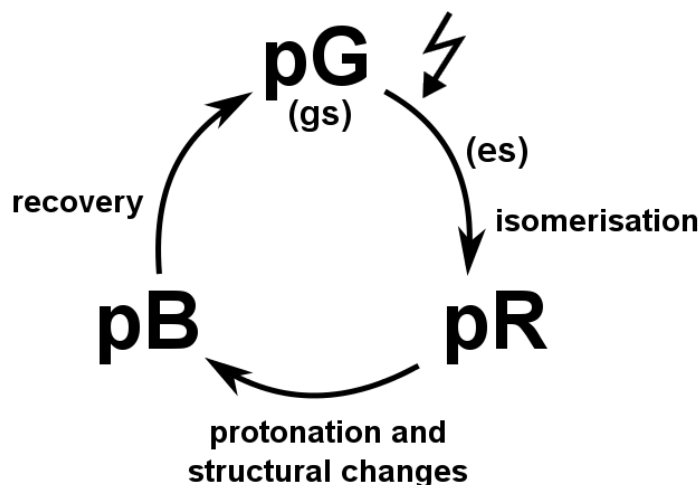


Figure 1.3: The PYP photocycle with the ground state (gs, pG), the excited state (es) after photon absorption by the chromophore, the pR state after deactivation of the chromophore from the excited state via db isomerisation, and the signalling state (pB), where protonation and structural changes are observed.

1.1.1 Chromophore Deactivation

In previous theoretical studies, the chromophore shows two possible excited state deactivation pathways: 1) Isomerisation of the double bond (db) and 2) rotation around the single bond (sb) adjacent to the phenyl ring. The 90 degree twisted structures of both pathways are depicted in Figure 1.4. These studies present deactivation channels in different media. In vacuum, the chromophore shows only deactivation via the phenyl ring rotation because this pathway has, in contrast to the isomerisation, no barrier to be overcome^{25–28}. In water, hydrogen bonds to the surrounding water molecules lower the barrier for the db isomerisation pathway and for the db twisted minima on the first excited state potential energy surface (PES). For deactivation via sb rotation, as in vacuum, no barrier exists between the Franck-Condon-region and the sb twisted minimum. Together with the fact that the conical intersection between the first excited state and the ground state is close to both minima, fast radiationless decay becomes possible. Because of the absent energy barrier for sb rotation and the existing energy barrier for db isomerisation, the chromophore is still preferentially deactivated via the pathway involving phenyl ring rotation^{26–29}.

So far, deactivation via sb rotation could not be observed convincingly in experiment because the resulting ground state structure after a 180 degree sb rotation can not be distinguished from the original ground state structure. Observations of an intermediate which does not lead to a productive photocycle but back to the ground state could thus originate from deactivation via sb rotation or from unsuccessful deactivation via db isomerisation. The comparison of theoretical and experimental data does not contradict the existence of this deactivation pathway and experimental studies consider deactivation via sb rotation during data analysis. Barrierless deactivation to *cis*-conformation in the ground state via db isomerisation has been observed in experimental studies on the chromophore models *p*-hydroxycinnamate pCA^{2–} and its amide analogue pCM[–] by Chagenet-Barret *et al.*^{30–32}. When embedded into the protein, the chromophore is surrounded by various amino acids that form consistent hydrogen bonds, a stable electrostatic environment and result in steric constraints on the chromophore conformation. As in water, theoretical studies observed that the hydrogen bonds to the chromophore lower the barrier of the db isomerisation pathway. Hence, the chromophore deactivates from the first excited state via the db isomerisation route^{27, 29, 33}. Experimental studies on PYP wild type and various mutants^{34, 35} show two competitive deactivation pathways. One of these leads to the productive photocycle with formation of the *cis*-isomer whereas the second is short lived and falls back to a *trans*-configured ground state intermediate. For both pathways deacti-

vation via db isomerisation is proposed.

The following mechanism was proposed based on theoretical studies^{26,36}. After photon absorption the negative charge, located mainly at the phenyl oxygen in the ground state, is translocated towards the ethylene chain of the chromophore. This translocation changes the properties of the chromophore by decreasing the double bond character of the C2-C3 bond for more flexibility^{37,38}, which allows the isomerisation of the ethylene chain whereas the phenyl group stays in place^{22,39}. This isomerisation happens via a concerted motion of the bonds belonging to the ethylene chain and subsequent breakage of the hydrogen bond to Cys69^{25,40}.

Premvardhan *et al.* measured the changes of the electrostatic properties between ground and excited state of the PYP chromophore by Stark spectroscopy⁴¹. The experiments showed a significant change in the permanent dipole moment upon photo excitation for both, a negatively charged chromophore model in water and the chromophore in PYP. The measured data is consistent with the movement over 5 Å of one electron or over a longer range for a partial charge. By theoretical calculations they assign this changes to a charge shift from the phenyl ring to the ethylene chain. This observation is in agreement with the aforementioned theoretical studies.

This thesis aims at gaining further information on the intrinsic deactivation pathways and the effect of different molecular environments on the deactivation by varying specific aspects. One important aspect of this environment is the hydrogen bonded network formed by the chromophore and the surrounding amino acids. To understand this aspect, the following section briefly summarises relevant information about hydrogen bonds.

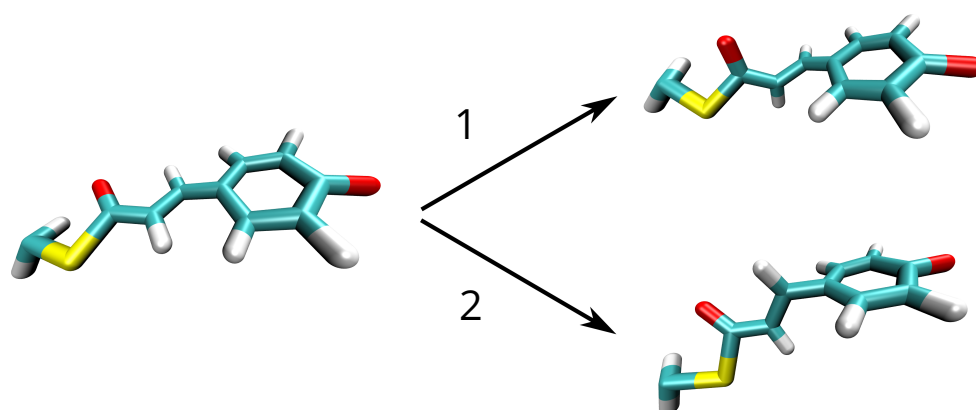


Figure 1.4: The planar chromophore in the ground state and the twisted structures resulting from the two possible deactivation pathways: double bond isomerisation (1) and single bond rotation (2)

1.2 Hydrogen Bonds

In bio-molecules, hydrogen bonds are important chemical interactions. They are responsible for the secondary structure of proteins, and DNA and for quaternary structures as the connection of two DNA strands to form the characteristic double helix. Hydrogen bonds can be both intra- (within the same molecule) or intermolecular (between two different molecules).

The PYP chromophore forms three hydrogen bonds to surrounding amino acids. Experimental studies on mutants, where these amino acids were exchanged to remove the hydrogen bonds, show alterations of the PYP photocycle^{19,34,35,42}. Additionally, two of these hydrogen bonds are exceptionally short, which will be described in more detail in chapter 3. Hydrogen bonds are of particular interest for this thesis because they influence the early part of the photocycle that includes the deactivation from the excited state. The theoretical concept of hydrogen bonds will be briefly outlined in the following paragraphs.

A hydrogen bond is an attractive interaction between a hydrogen donor (A) and a hydrogen acceptor (B) by which two electronegative groups interact via a proton (H). The hydrogen donor is formed by an electronegative group to which a proton is covalently bound. This proton is attracted to the hydrogen acceptor, which is an electronegative group with a free electron pair, but not strong enough to form a covalent bond. The interaction is electrostatic and can be described as a dipole-dipole interaction. It is weaker than covalent or ionic bonds. The donor-acceptor distance is typically 2.7–3.0 Å long and the hydrogen bond up to 16 kJ/mol strong^{43,44}.

There are several concepts for potential energy curves regarding the proton position which could characterise hydrogen bonds⁴⁵:

1. The potential energy curve for the hydrogen position of a typical hydrogen bond is an asymmetrical double well potential where the proton is attracted more strongly to the hydrogen donor than to the hydrogen acceptor. The proton is, as described above, covalently bound to the hydrogen donor. This is due to the fact that the two electronegative atoms are not identical in their chemical properties (Figure 1.5).
2. A symmetric double well potential with a barrier close to or lower than the zero point energy can be referred to as low barrier hydrogen bond (LBHB). In this case, the proton is able to be covalently bound to both

electronegative atoms and switch between those two states. Both states have not to be equally populated (Figure 1.6).

3. Single well potentials might also describe hydrogen bonds. When the donor-acceptor distance gradually approaches an unusual length of less than 2.5 \AA the barrier between the two minima disappears to form a single minimum potential energy curve. The proton is centered between donor and acceptor⁴⁶. For these hydrogen bonds, the term short strong hydrogen bond (SSHB) is used (Figure 1.7).

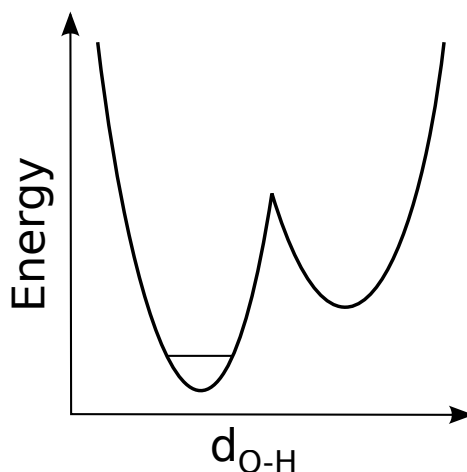


Figure 1.5: Asymmetrical hydrogen bond potential for a typical hydrogen bond. The two electronegative atoms have different chemical properties and the donor-acceptor distance is $2.7\text{--}3 \text{ \AA}$. The hydrogen is covalently bound to the hydrogen donor..

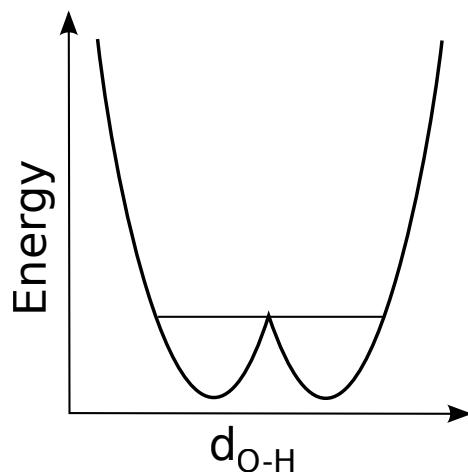


Figure 1.6: A low barrier hydrogen bond (LBHB) potential has two equal or almost equal minima along the hydrogen bond. The barrier separating them is low enough for the hydrogen to occupy both forming covalent bonds to both donor and acceptor. This symmetrical double well potential arises from similar electronegativity of hydrogen donor and acceptor and shorter donor-acceptor distance.

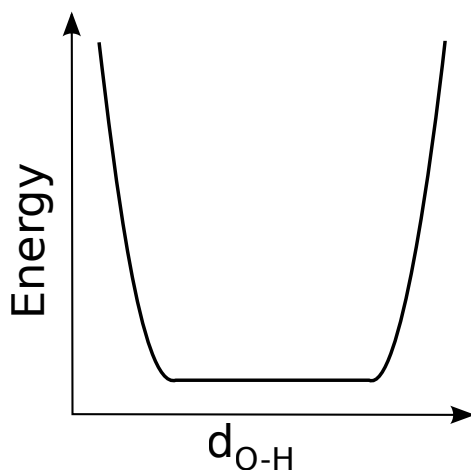


Figure 1.7: In a single well hydrogen bond potential no energy barrier is present along the bond between hydrogen donor and acceptor. The hydrogen is centred between both electronegative atoms. Donor and acceptor have almost the same electronegativity and the distance between them is less than 2.5 Å.

The concept of the third type of hydrogen bonds was introduced because hydrogen bonds shorter than 2.55 Å were discovered in X-ray structures of several proteins, indicating the presence of a SSHB. Furthermore, neutron diffraction studies showed the proton to be in the centre of the bond at equal distance between donor and acceptor^{47,48}, also supporting the existence of a SSHB. For PYP, this was observed by Yamaguchi *et al.*⁴⁹. A detailed description of their observations can be found in chapter 3. The role of stabilising intermediates or transition states was assigned to SSHBs in proteins⁴⁵. These hydrogen bonds were described to have a covalent contribution to the bond or even being completely covalent (4 electron, 3 center bonds)⁵⁰. For SSHBs in general striking properties were predicted^{51–54}:

- The strength is up to 160 kJ/mol.
- The energy potential along the bond is a symmetrical single well.
- The A-H distance increases for decreasing A-B distance.
- H has equal distance of 1.2 Å to both A and B at an A-B distance of 2.4 Å.
- pK_a of A and B is almost identical.

There are mainly two criticisms for the concept of SSHBs. First, calculations show that the reduction of the donor-acceptor distance stabilises the bond by just 16 kJ/mol, whereas the position of the proton in the middle of the bond destabilises it by about 60 kJ/mol⁵⁵. Second, SSHBs could only be observed for the FHF[−] molecule in gas phase and some solid state molecules⁵⁶. All proteins in solution studied by Perrin *et al.*⁵⁵ did not contain SSHBs. Notably shorter hydrogen bonds could be found but all other above described properties could not be verified for proteins. The short hydrogen bonds found in Perrin’s study always show a slight asymmetrical potential. These systems are described as low barrier hydrogen bonds (LBHBs)⁵⁵. Overall, short and apparently strong hydrogen bonds could be found for charged donors and acceptors in gas phase or solid state. All postulated SSHBs (in proteins or in solution) other than FHF[−] in gas phase and some molecules in solid state have been disproved.

The observations can be explained by the fact that the localised charge of an asymmetric hydrogen bond is stabilised more in polar environments than the delocalised charge of a symmetric hydrogen bond⁵⁷. Thus, polar residues in an enzyme close to an assumed SSHB would destabilise it⁵⁸. In contrast to the presumption that short hydrogen bonds are especially

strong, Perrin *et. al.*⁵⁵ concluded from their study that short hydrogen bonds are strained by external factors, for example the fixed geometry in a crystallised protein, that force the donor atoms together. This strain results in a destabilisation of the hydrogen bond in contrast to the presumed stabilisation due to an extraordinary short hydrogen bond. The strain would be relieved by elongation of the hydrogen bond which would be energetically favourable. According to this argument, SSHBs can be found in crystals, where molecular structures are compressed, but not in solution.

1.3 Motivation

Photochemical reactions are important for biological functions in many aspects. Photosynthesis is the main energy source for plants by producing glucose from carbon dioxide and water using solar energy. Vitamin D is synthesised by humans using a photochemical reaction. Isomerisations of double bonds initiated by light absorption are essential for the mechanism of vision. The retinal chromophore in the rhodopsin protein isomerises upon light absorption and the subsequent conformational changes trigger a signalling cascade which enables us to see.

In PYP the isomerisation after light absorption is the first step of a signalling cascade as well. It leads to a negative phototactic response of the bacterium *Halorhodospira halophila*.

Possible ways to influence the PYP photocycle by controlling the chromophore deactivation are of strong interest because PYP is the model protein for the PAS domain super family which consists of signal transduction proteins. This question is in particular relevant with the observed dependency of the chromophore deactivation with respect to its environment. By putting the chromophore first from vacuum into solution and then into the protein, its deactivation process upon light excitation changes from the rotation around the single bond to the isomerisation of the double bond with simultaneous movements of adjacent bonds²⁷. One point of interest is how the various amino acids of the binding pocket affect this process.

The enclosure of the chromophore inside the binding pocket may disable the deactivation pathway via rotation of the phenyl ring due to sterical hindrance. Thus, the isomerisation of the ethylene chain, which is *trans*-configured in the ground state, is the favoured deactivation pathway of the protein bound chromophore from the first excited state.

X-ray structures of PYP show three hydrogen bonds formed by the chromophore to Tyr42, Glu46 and Cys69. Two hydrogen bonds have been found to be in the length range of the above described SSHBs. For example, in a structure recently published by Yamaguchi *et al.*⁴⁹, hydrogen bond lengths of 2.52 and 2.56 Å were found from the chromophore to Tyr42 and Glu46, respectively. For the hydrogen bond between the chromophore and Glu46, the proton was observed to be in the middle of the hydrogen bond, which is consistent with the SSHB theory. In chapter 3, this thesis will answer the following question: Can the experimentally observed SSHB be reproduced by theoretical methods and which conditions are required?

The answer to this question can help to understand the terms under which

an SSHB can occur in a protein and provide a better insight into the role of hydrogen bonds in the PYP photocycle and thereby in other photocycles. To address this question, a series of QMMM single point calculations was applied. To control the parametrisation of the model, the number of residues that were considered by the QM method has been varied. As a result, insights into the required environmental specifications for modelling the SSHB were gained.

Chapter 4 will deal with the protonation state of residue Arg52, which is one of the amino acids forming the chromophore binding pocket. This residue has been found to be deprotonated in the above mentioned study by Yamaguchi *et al.* The amino acid arginine is protonated in most proteins for which reason the experimental observation leads to the question: How does the protonation state of Arg52, which separates the chromophore binding pocket from bulk water, influence the protein dynamics and deactivation events?

To answer this question MM and QMMM MD simulations were carried out and were performed with corresponding setups for both protonated and neutral Arg52.

The influence of the hydrogen bond network on the chromophore deactivation from the first excited state is topic of chapter 5. Two solvents with different viscosity and polarity, water and decanol, were used to solvate the chromophore model pCK⁻. QMMM MD simulations were applied to monitor the deactivation events of the chromophore. With these calculations the following question was addressed: How does the hydrogen bonded network around the chromophore affect the deactivation from the first excited state?

All three questions are related to the environment around the chromophore, especially the hydrogen bonds formed by the chromophore's phenyl oxygen, and their influence on the deactivation from the excited state. This thesis aims at giving new insights into this matter by applying theoretical methods to a chromophore model and the protein PYP.

The simulations were carried out using the programs Gromacs⁵⁹, Gaussian^{60,61}, and MolPro⁶². For visualisations the program VMD^{63,64} was used.

Chapter 2

Methods & Theory

To study the light induced processes in PYP in atomistic detail, suitable computational methods have to be used. In this work, standard methods as described in textbooks (Frank Jensen⁶⁵, Christopher J. Cramer⁶⁶ and the collection of the lecture notes of the ESQC2011⁶⁷) were used. Relevant details are described in the following sections.

For the following descriptions, a system is defined as investigated molecule(s) with respective environment (solvent, ions etc.) and conditions (temperature, pressure etc.).

As a biological molecule is flexible, Molecular Dynamics is a valuable method and used for most calculations in this thesis.

2.1 Molecular Dynamics

Molecular Dynamics (MD) is a computational method to calculate ensemble averages of the system by following the time course^{66,68,69}.

For proteins an X-ray structure of the desired molecule is used as starting structure for MD simulations but for small molecules the structures can also be build with suitable computer programs. Before working with a crystal structure, it should be checked if it is consistent with known data of the protein and if the free R-factor, which is a value for the cross-validation in the refinement process, is sufficiently low. The X-ray structure is refined by adding missing atoms and minimising the whole structure via Molecular Mechanics (see section 2.2) to reduce tensions. These tensions are caused by changes made to the crystal structure like adding hydrogen atoms or other missing residues that were not resolved by X-ray crystallography.

To calculate the dynamical properties, initial velocities are assigned to every atom, based on the Maxwell-Boltzmann distribution for a given temperature. For every time step during an MD simulation, the positions and velocities of each atom are calculated by integrating Newton's equation in finite time steps. The Verlet algorithm⁷⁰ is one of the simplest but often the appropriate algorithm for this task.

When a forward and backward Taylor series expansion of the coordinate $r(t)$ of a particle at time t is carried out:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{a(t)}{2}\Delta t^2 + \frac{b(t)}{3!}\Delta t^3 + \mathcal{O}(\Delta t^4) + \dots \quad (2.1)$$

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{a(t)}{2}\Delta t^2 - \frac{b(t)}{3!}\Delta t^3 + \mathcal{O}(\Delta t^4) + \dots, \quad (2.2)$$

the velocity $v(t)$ is the first derivative of r with respect to the time, $a(t)$ the second derivative, and $b(t)$ the third derivative.

Summing both equations gives

$$r(t + \Delta t) + r(t - \Delta t) = 2r(t) + a(t)\Delta t^2 + \mathcal{O}(\Delta t^4). \quad (2.3)$$

By neglecting the last term new positions are calculated with an error of order Δt^4 , Δt being the time step of the MD simulation.

Therefore, velocities are not needed to calculate the new positions, but can be determined from the positions by applying

$$r(t + \Delta t) - r(t - \Delta t) = 2v(t)\Delta t + \mathcal{O}(\Delta t^3) \quad (2.4)$$

giving

$$v(t) = \frac{r(t + \Delta t) - r(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2). \quad (2.5)$$

The newly calculated velocities have an error of order Δt^2 . To get more accurate results for the velocities, other Verlet-like algorithms, e.g., the leap-frog-algorithm, are used in this work^{71,72}.

The leap-frog-algorithm has the same order of integration error as the Verlet algorithm but performs better in terms of numerical accuracy. Its derivation is similar to the Verlet algorithm but the positions and velocities are derived with half a time step in between, giving the following equations:

$$r(t + \Delta t) = r(t) + v(t + \frac{1}{2}\Delta t)\Delta t \quad (2.6)$$

and

$$v(t + \frac{1}{2}\Delta t) = v(t - \frac{1}{2}\Delta t) + a(t)\Delta t. \quad (2.7)$$

As the velocities are calculated, the leap-frog-algorithm supports the coupling of the system to an external heat bath, which allows to keep the temperature constant. The temperature regulation is done by adjusting the velocities to represent the Maxwell-Boltzmann distribution for the assigned temperature.

The easiest way for the temperature regulation would be to determine the temperature at each time step, compare it to the desired temperature and then scale all velocities to get the desired temperature, accordingly. This method produces too much disturbances and the systems evolution is biased. In this work, the Berendsen thermostat⁷³ and v-rescale⁷⁴ are used. With the Berendsen thermostat, the velocities are not simply scaled to the desired temperature causing a temperature jump, but are smoothly adjusted by the use of an integrational time step and a scaling constant. The time it takes to adjust the temperature to the desired value after a temperature change occurred depends on the size of the scaling constant.

V-rescale is another approach to adjust the temperature of the system. Here the actual kinetic energy (E_T^a) and a target value, which is calculated from the equilibrium distribution of the kinetic energy at the target temperature (E_T^t), are consolidated into a scaling factor

$$s = \sqrt{\frac{E_T^t}{E_T^a}}. \quad (2.8)$$

This factor is applied to all atoms similar to the Berendsen thermostat to avoid considerable disturbances of the system⁷⁴.

MD simulations can be carried out using Molecular Mechanics as well as Quantum Mechanics to describe interactions between atoms. Molecular Mechanics uses various approximations which is computationally efficient and hence used for modelling most of the system in MD simulations in this work and will be described in the subsequent section.

2.2 Molecular Mechanics

In Molecular Mechanics (MM) parameters and functions are comprised in a force field, to calculate the properties of a system applying classical mechanics.

The fundamental equation of MM is

$$E = E_{\text{bonded}} + E_{\text{non-bonded}}, \quad (2.9)$$

where the summands consist of the following:

$$E_{\text{bonded}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}$$

and

$$E_{\text{non-bonded}} = E_{\text{electrostatic}} + E_{\text{van der Waals}}.$$

Most of the computational cost in MM simulations is due to non-bonded terms, because of their large range, which results in interactions between all atoms, whereas the bonded interactions apply only to the next atoms.

The bonding and angle terms are generally calculated with harmonic potentials around the equilibrium values and consist of sums over these respective terms

$$E_{\text{bond}}(r) = \frac{1}{2}k_r(r - r_0)^2 \quad (2.10)$$

and

$$E_{\text{angle}}(\theta) = \frac{1}{2}k_\theta(\theta - \theta_0)^2. \quad (2.11)$$

Dihedrals have more than one minimum and can not be calculated harmonically, but are computed using different functions depending on the program, applying

$$E_{\text{dihedral}}(\omega) = \sum_n V_n \cos(n\omega). \quad (2.12)$$

The van der Waals term declines rapidly and is often modelled using the 6-12-Lennard-Jones-potential⁷⁵, where the attractive forces decline with r^{-6} and the repulsive forces decline with r^{-12} :

$$E_{\text{van der Waals}} = E_{\text{LJ}}(r) = \sum_{ij} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (2.13)$$

with the well depth of the potential ε , and the distance at which the interparticle potential is zero σ .

To speed up the calculations, a cut-off-radius is applied, which assists the calculation of van der Waals interactions for atoms that are far apart.

Electrostatic terms do not decline as rapidly as van der Waals terms. These electrostatic interactions are therefore important for the adequate description of proteins with MM. They are given by the Coulomb-potential:

$$E_{\text{electrostatic}} = E_C(r) = \frac{1}{4\pi\varepsilon_0} \sum_{ij} \frac{q_i q_j}{r_{ij}}. \quad (2.14)$$

To limit the calculation time, also here, a cut-off radius is applied. This cut-off results in a considerable discontinuity at the border between the region where the electrostatics are taken into account and the region where these are neglected. By using switching or scaling functions, this discontinuity is tried to be flattened and to resemble the results that are obtained when the electrostatic energy is taken into account without cut-offs.

Summing up all terms, of which the fundamental equation of MM (Equation (2.9)) consists, yields the complete MM energy:

$$\begin{aligned} E = & \sum_{\text{bonds}} \frac{1}{2} k_r (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{\text{torsions}} \sum_n V_n \cos(n\omega) \\ & + \sum_{ij} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\varepsilon_0} \sum_{ij} \frac{q_i q_j}{r_{ij}}. \end{aligned} \quad (2.15)$$

One application of MM is energy optimisation. For this purpose, parameters from a force field are used. The local minimum of the energy next to the starting conformation is found using a suitable optimisation algorithm. The simplest way to find a local minimum is the steepest descent method (SD). The force calculated for a system that is not in equilibrium is the negative of the energy gradient. Following the force until it is zero moves the system towards the local minimum. This method leads quickly to improvements when the gradient is large, but is very slow for small gradients. These are often found near the local minimum. To accelerate the determination of the local minima, the Newton-Raphson algorithm is used. This algorithm requires the second derivative of the energy, the Hessian. As calculating the Hessian is computationally more expensive minimisations, are started using the SD method and near the minimum, when convergence is slow, the Newton-Raphson method is used.

2.3 Hybrid Quantum Mechanical/Molecular Mechanics

For detailed analysis in the context of the whole protein, hybrid Quantum Mechanics/Molecular Mechanics models have been developed and will be described in this section.

Quantum Mechanics (QM) is a very accurate description for a system. But only very small systems can be modeled with *ab initio* QM methods. Even with semi-empirical methods, simulations of whole proteins are very time-consuming. Instead, MM can be used to calculate large systems like proteins, but all bonds, angles, and dihedrals come from a parametrized force field. Because of this, it is impossible to model polarisation, electron transfer, or bond breaking and formation, which is possible with QM methods. To combine the advantages of both methods while the disadvantages are kept minimal, a hybrid of both approaches is used in this work.

Specifically, the part of the molecule that is of research interest is treated with QM whereas the remaining system is treated with MM as demonstrated in Figure 2.1^{66,76,77}.

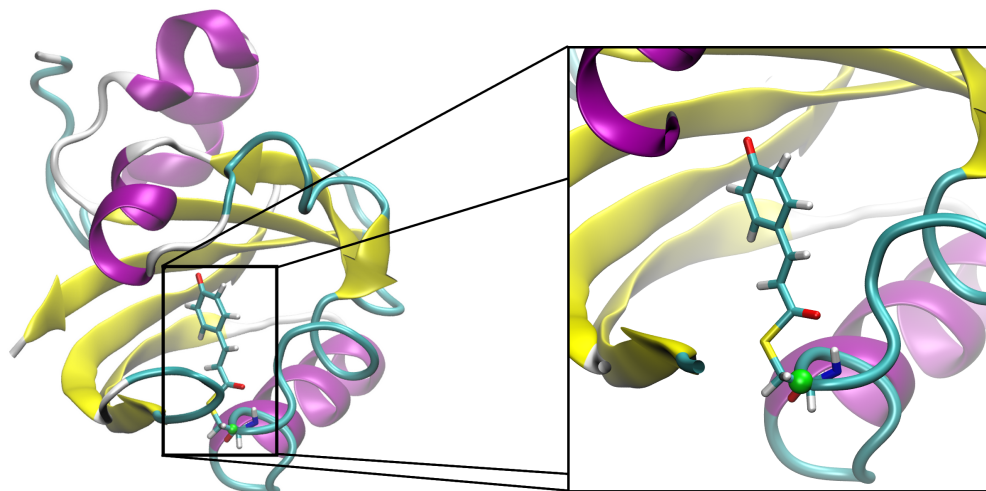


Figure 2.1: PYP with the chromophore as the QM region (depicted as explicit atoms) and all other atoms forming the MM part (depicted using the cartoon scheme). Both regions are connected by the link atom (depicted as green sphere), which saturates the QM region and gives the MM parameters an atom to act upon.

The effective Hamiltonian for a QMMM system,

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{QM/QM}} + \hat{H}_{\text{QM/MM}} + \hat{H}_{\text{MM/MM}}, \quad (2.16)$$

consists of $\hat{H}_{\text{QM/QM}}$, the Hamiltonian of the isolated QM-part, $\hat{H}_{\text{MM/MM}}$, the Hamiltonian of the isolated MM-part, and $\hat{H}_{\text{QM/MM}}$, which describes the interactions between both parts.

The energy expectation value for a state $|\Phi\rangle$ is given by

$$E = \langle \Phi | \hat{H}_{\text{eff}} | \Phi \rangle = \langle \Phi | \hat{H}_{\text{QM/QM}} | \Phi \rangle + \langle \Phi | \hat{H}_{\text{QM/MM}} | \Phi \rangle + E_{\text{MM/MM}}. \quad (2.17)$$

The QM/MM Hamiltonian is modelled by electrostatic and van der Waals interactions

$$\begin{aligned} \hat{H}_{\text{QM/MM}} = & \sum_i^{\text{QM electrons}} \sum_m^{\text{MM atoms}} \frac{q_m}{r_{im}} \\ & + \sum_k^{\text{QM atoms}} \sum_m^{\text{MM atoms}} \left[\frac{Z_k q_m}{r_{km}} + 4\varepsilon_{km} \left(\frac{\sigma_{km}^{12}}{r_{km}^{12}} - \frac{\sigma_{km}^6}{r_{km}^6} \right) \right], \end{aligned} \quad (2.18)$$

with the counters i for the number of electrons, k for the number of nuclei, and m for the number of MM atoms. The other variables are q_m for the atomic charge on MM atoms, Z for the nuclear charge, r for the distances between particles, and ε and σ for the well-depth and equilibrium distance of the Lennard-Jones potential.

With this hybrid approach, one has to define how the molecular parts in the two regions interact. For systems where no bonds are part of both regions, no special action has to be taken. For systems with bonds between both regions, two requirements have to be fulfilled. First, the QM-atoms have to be saturated and second, the bond-, angle-, and dihedral terms of the force field for the atoms adjacent to the bonds need a partner to act upon. To solve this problem, a pseudo hydrogen atom, the link atom, is inserted into the bond connecting the QM and MM region. This link atom allows the valence of the last QM-atom before the MM-region to be saturated and the MM-terms to have an atom referencing the geometry of the system.

For this work, two methods were used to describe the QM region: Complete Active Space Self Consistent Field and Density Functional Theory, which are described in more detail in the following sections. For both methods, two approximations are applied:

- The electronic Hamiltonian depends only on the positions of the nuclei, but is independent of their respective momentum. As the electrons move much faster than the nuclei because of their low mass, the coupling between the electronic and the nuclei motion can be neglected. This approach is called the Born-Oppenheimer-Approximation, which says that the electrons move in the field of fixed nuclei.
- As the relativistic effects are not important for calculations with elements of the first three rows of the periodic table, unless high accuracy is needed, these are neglected.

2.4 Density Functional Theory

To specify the concept of Density Functional Theory (DFT) and Complete Active Space Self Consistent Field (CASSCF) the Hartree-Fock approach will be briefly introduced.

Hartree-Fock

Hartree-Fock is a quantum mechanical method for calculating wave function and energy of the ground state of a many-body system by solving the time-independent Schrödinger equation. Several approximations are used to get a computationally efficient method that gives adequate results at the same time.

- The wave function of each electron is described as an orbital, given by a linear combination of basis functions out of a preassigned set. The product of those single electron orbitals forms the many particle basis for the total electron wave function. The Pauli principle excludes that two electrons share the same orbital. Thus the wave function has to be totally anti-symmetric which means that it changes the sign when any two electrons are exchanged. This totally anti-symmetric product is called Slater determinant.
- By using the variational principle, the orbitals are optimised to give the lowest possible energy for the chosen basis set, restricting the total wave function to be described by a single Slater Determinant.
- The mean field approximation used in Hartree-Fock averages the interactions between the electrons. A single electron does not interact with each other electron separately, but with an averaged electron cloud.

In contrast to the wave function used in the Hartree-Fock- and post-Hartree-Fock-methods, as CASSCF, which will be described in the following section, the electron density $\rho(r)$ is a physical observable, which only depends on \vec{r} . Density Functional Theory (DFT) derives all physical quantities from the electron density $\rho(r)$. This is given for a N-electron wave function Ψ by

$$\rho(r_1) = N \int |\Psi(x_1, x_2, \dots x_N)|^2 dx_2 \dots dx_N. \quad (2.19)$$

Hohenberg and Kohn (HK)⁷⁸ demonstrated that

- the ground state electron density is uniquely defined by the core potential and the ground state electron density uniquely defines the core potential (HK theorem I).
- the variational principle holds for calculating the energy from the electron density. The ground state electron density minimises the energy (HK theorem II).

From the HK theorem I, it follows that the energy is given by a functional of the electron density $E[\rho]$ split into parts as

$$E[\rho] = T[\rho] + E_{ne}[\rho] + E_{ee}[\rho]. \quad (2.20)$$

The term $E_{ne}[\rho]$ for the attraction between electrons and nuclei is given by the expression

$$E_{ne}[\rho] = - \sum_a^{N_{nuclei}} \int \frac{Z_a(\mathbf{R}_a)\rho(\mathbf{r})}{|\mathbf{R}_a - \mathbf{r}|} d\mathbf{r}. \quad (2.21)$$

In equation 2.20, the electron-electron repulsion $E_{ee}[\rho]$ is given by the Coulomb $J[\rho]$ and exchange parts $K[\rho]$. The Coulomb energy term is

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (2.22)$$

Terms for the kinetic and exchange energy are approximated, e.g., from the electron density of a free electron gas. In these approximations the exchange and kinetic terms are given by

$$K[\rho] \simeq -C_x \int \rho^{\frac{4}{3}}(\mathbf{r}) d\mathbf{r}$$

and

$$T[\rho] \simeq C_T \int \rho^{\frac{5}{3}}(\mathbf{r}) d\mathbf{r}, \quad (2.23)$$

where C_x and C_T are constants.

A solution to this problem was presented by Kohn and Sham (KS)⁷⁹. They proposed to reintroduce orbitals and for this split the kinetic energy

into one that can be calculated exactly for a non-interacting system and a small correction to the exact term that is absorbed into the exchange-correlation term. By this approach the kinetic energy becomes equivalent to the Hartree-Fock kinetic energy

$$T_S = \sum_{i=1}^{N_{elec}} \langle \phi_i | -\frac{1}{2} \nabla^2 | \phi_i \rangle. \quad (2.24)$$

The total energy for DFT can thus be formulated as

$$E_{DFT}[\rho] = T_S[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho]. \quad (2.25)$$

By comparing the DFT energy (Equation 2.25) to the exact energy calculated from the electron density (Equation 2.20), the exchange-correlation term is defined as

$$E_{xc}[\rho] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]). \quad (2.26)$$

Exchange-Correlation Functionals

Different DFT methods vary by the exchange-correlation functional used. A correct exchange-correlation functional is only possible for a uniform electron gas which is used in the first method mentioned below. For all other systems, the functional is an approximation.

The popular types of Exchange-Correlation Functionals are the following:

1. The local density approach (LDA) expresses the energy as the integral over a function of the density

$$E_{xc} = \int F(\rho) d\mathbf{r}. \quad (2.27)$$

2. The generalised gradient approximation (GGA) expresses the energy as the integral of a function of the density and its gradient

$$E_{xc} = \int F(\rho, \nabla \rho) d\mathbf{r}. \quad (2.28)$$

3. The meta GGAs use higher order derivatives of the density

$$E_{xc} = \int F(\rho, \nabla \rho, \nabla^2 \rho) d\mathbf{r}. \quad (2.29)$$

4. The hybrid functionals combine a GGA functional with a portion of exchange energy calculated from orbitals

$$E_{xc} = \int F(\rho, \nabla \rho) d\mathbf{r} + \xi E_X^0. \quad (2.30)$$

The most established exchange-correlation functional is the hybrid functional B3LYP, which is also used in this work. It is a functional of type 4 with $\xi = 20\%$. This functional's popularity is because of its good overall performance.

Time-dependent DFT (TD-DFT) is, in contrast to DFT, able to describe excited states that are important for the investigation of the deactivation events in PYP, the topic of this work, and has been used for excited state MD simulations⁸⁰⁻⁸².

DFT as well as Hartree-Fock are single-reference methods. These methods are not able to describe bond breaking and formation sufficiently, as will be described in the following section on the example of breaking the bond in H₂ using Hartree-Fock.

Further problems encountered by TD-DFT are the description of valence states in molecules with extended π -systems, charge-transfer excited states⁸³ and conical intersections between ground and excited state^{84,85}. Both an extended π -system as well as conical intersections are present in the deactivation events in PYP studied in this thesis. Therefore TD-DFT will not be used. For the excited state simulations CASSCF, described in the following section, is applied.

As DFT, in particular the B3LYP functional, is a reliable and computationally efficient QM method, it is used in this work for ground state calculations.

2.5 Complete Active Space Self Consistent Field (CASSCF)

CASSCF is a multiconfigurational post-Hartree-Fock-method and thus overcomes the restrictions of Hartree-Fock in the case of e.g., bond breaking or excited state calculations. For systems with nearly degenerate states, such as a ground state and a low lying excited state, it is not sufficient to use one Slater Determinant as it is done in the Hartree-Fock method.

Going beyond the Hartree-Fock approach is illustrated by the example of the H_2 molecule. This shows both the need for post-Hartree-Fock-methods and the principle of the problems with Hartree-Fock.

The simple H_2 -molecule can not be described adequately by restricted Hartree-Fock (RHF). Whereas the molecule is described sufficiently at the equilibrium geometry, its dissociation is not described accurately, which is seen by the following consideration. The molecular orbitals of H_2 can be written as

$$\varphi_i = N_i(\chi_{iA} \pm \chi_{iB}). \quad (2.31)$$

The doubly occupied molecular orbital φ_1 is the bonding orbital built from the atomic orbitals χ_{1A} and χ_{1B} . As the molecular orbitals are formed in a minimal basis from the 1s atomic orbitals, the orbital can be written as

$$\varphi_1 = N_1(1s_A + 1s_B). \quad (2.32)$$

The RHF model assumes the bonding molecular orbital to be doubly occupied what leads to the total wave function

$$\Psi_1 = \varphi_1(\mathbf{r}_1)\varphi_1(\mathbf{r}_2)\Theta_{2,0}, \quad (2.33)$$

where $\Theta_{2,0}$ is the singlet spin function ($S = 0$) for two electrons

$$\Theta_{2,0} = \sqrt{\frac{1}{2}}(\alpha_1\beta_2 - \beta_1\alpha_2). \quad (2.34)$$

The RHF model gives rise to large errors in the molecule's behaviour due to its strong distance dependency. These errors happen because the electron correlation energy is missing even if it is just a small fraction of the total energy because in RHF the individual electron interacts with an average static electric field created by all other electrons. The electron correlation energy is the energy missed by the RHF approach. It corresponds to the correlated motion of the electrons.

The description of the MO (Equation 2.32) is distance independent. Expanding this wave function yields

$$\begin{aligned}\Phi_1 = & N_1^2 [1s_A(\mathbf{r}_1)1s_A(\mathbf{r}_2) + 1s_A(\mathbf{r}_1)1s_B(\mathbf{r}_2) \\ & + 1s_B(\mathbf{r}_1)1s_A(\mathbf{r}_2) + 1s_B(\mathbf{r}_1)1s_B(\mathbf{r}_2)]\Theta_{2,0}.\end{aligned}\quad (2.35)$$

This equation shows that the wave function contains ionic terms where both electrons can be found at one atom. At large distances, these terms are unphysical as they represent the state $H^+ + H^-$ and give rise to wrong dissociation energies. As ionic terms are weighted equally in the RHF model, it can not describe dissociation processes where open shells are present in the products.

A solution is the introduction of coefficients for the different terms in the wave function, which gives rise to the valence bond formulation of the wave function

$$\Psi_{VB} = C_{Ion}\Psi_{Ion} + C_{Cov}\Psi_{Cov}.\quad (2.36)$$

To yield the correct wave function, the coefficients C_{Ion} and C_{Cov} can be varied to $C_{Ion} \approx 0$ for the dissociated molecule and to $C_{Ion} \approx C_{Cov}$ for the molecule at the equilibrium geometry.

Equation 2.36 is formulated in terms of non-orthogonal basis functions

$$\Psi_{Ion} = N_{Ion} [1s_A(\mathbf{r}_1)1s_A(\mathbf{r}_2) + 1s_B(\mathbf{r}_1)1s_B(\mathbf{r}_2)]\Theta_{2,0}$$

and

$$\Psi_{Cov} = N_{Cov} [1s_A(\mathbf{r}_1)1s_B(\mathbf{r}_2) + 1s_B(\mathbf{r}_1)1s_A(\mathbf{r}_2)]\Theta_{2,0}.\quad (2.37)$$

An equivalent formulation is yielded by the introduction of the anti-bonding orbital

$$\varphi_2 = N_2(1s_A - 1s_B).\quad (2.38)$$

The expanded wave function for the anti-bonding orbital

$$\begin{aligned}\Phi_2 = & N_1^2 [1s_A(\mathbf{r}_1)1s_A(\mathbf{r}_2) - 1s_A(\mathbf{r}_1)1s_B(\mathbf{r}_2) \\ & - 1s_B(\mathbf{r}_1)1s_A(\mathbf{r}_2) + 1s_B(\mathbf{r}_1)1s_B(\mathbf{r}_2)]\Theta_{2,0}\end{aligned}\quad (2.39)$$

differs from the wave function for the bonding orbital in the sign in front of the ionic terms.

The wave function can now be written as the multiconfigurational (MC) molecular orbital formulation

$$\Psi_{MC} = C_1\Psi_1 + C_2\Psi_2, \quad (2.40)$$

which accounts for both the bonding and the antibonding orbital, and thus allows for excitations and is not restricted to doubly or unoccupied orbitals. This formulation is able to describe the complete potential curve more accurately. The coefficients are $C_1 \approx -C_2$ for the dissociated molecule and $C_1 \approx 1$ and $C_2 \approx 0$ for the molecule at equilibrium geometry.

Configuration Interaction

Adjusting the coefficients in front of the Slater determinants enables to account for all of the electron-electron interaction energy, which is only taken into account as averaged electron-electron interactions in the Hartree-Fock method. The missing part is called Electron Correlation (EC) energy. Electron correlation is defined as energy difference between the exact energy and the restricted HF energy in a given basis set

$$E_{EC}^{basis} = E_{exact}^{basis} - E_{RHF}^{basis}. \quad (2.41)$$

A way to take the electron correlation into account is to go from a one-determinant wave function to a method that uses more than one Slater determinant Φ ,

$$\Psi = a_0\Phi_{HF} + \sum_{i=1} a_i\Phi_i. \quad (2.42)$$

The HF Slater determinant for closed shells is constructed from $\frac{1}{2}N_{elec}$ molecular orbitals, which are multiplied by the two spin functions, to give N_{elec} spin orbitals. By promoting electrons from occupied spin orbitals into unoccupied virtual orbitals, several Slater determinants can be constructed. The number of possible Slater determinants is restricted by the number of virtual orbitals, i.e., the size of the used basis set. In accordance with the number of lifted electrons, these Slater determinants are singly (S), doubly (D), triply (T), etc. excited compared with the reference HF Slater determinant.

To find an electronic ground state, Configuration Interaction (CI) uses the variational principal to minimise the energy expectation value

$$\frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} \quad (2.43)$$

of a linear combination of determinants by varying the expansion coefficients a_i , minimising the energy Eigen value $\langle \Phi_{CI} | \hat{H} | \Phi_{CI} \rangle$.

The wave function

$$\Psi_{CI} = a_0 \Phi_{HF} + \sum_S a_S \Phi_S + \sum_D a_D \Phi_D + \sum_T a_T \Phi_T + \cdots = \sum_{i=0} a_i \Phi_i \quad (2.44)$$

includes all electron correlation effects and is called Full CI wave function. This applies for a given basis set when all possible excitations are taken into account. Therefore, the quality of calculations critically depends on the selected basis set.

Full CI is only possible to be solved for very small systems and basis sets. Hence it is needed to truncate the excitation level to get a computationally efficient method. As the matrix element $\langle \Phi_1 | \hat{H} | \Phi_2 \rangle$ of singly-excited determinants with the non-excited HF determinant is zero, using just these determinants does not give a better description of the ground state than pure HF. On that account, including exclusively doubly-excited determinants (CID) is the lowest level of CI at which the electron correlation contributes to the ground state. Adding singly-excited determinants to this gives the CISD method where the singly-excited determinants improve the doubly-excited determinants and, as a result, indirectly the ground state. CISD is the only computationally applicable method for medium sized systems and basis sets. Adding further excitations (as in CISDT or CISDTQ) would improve the recovered electron correlation, but would not be feasible for the system in this work.

The computationally needed restriction of the number of included determinants makes it necessary to add other techniques to describe certain systems sufficiently. Describing the wave function with more than one configuration is a technique to achieve a sufficient description.

Multi-Configurational Self-Consistent Field (MCSCF)

To go from truncated CI to Multi-Configuration Self-Consistent Field (MCSCF), not only coefficients in front of the determinants are optimised but also molecular orbitals that were used to construct the determinants.

Complete Active Space Self Consistent Field (CASSCF) is the most widely used MCSCF method and is illustrated in Figure 2.2. It utilises the fact that for the low excited states the occupation of orbitals differs just for the highest occupied and the lowest unoccupied orbitals of the ground state. These orbitals, in case of PYP six occupied and five unoccupied π -orbitals, are combined to form the so-called active space. All orbitals that are energetically lower than the active space remain fully occupied, whereas all orbitals with higher energy than the active space remain unoccupied. Within the active space, a Full CI calculation is carried out, whereas the other orbitals are relaxed on the SCF basis.

For calculations where more than one state is surveyed, state-averaging is employed. This method does not optimise the orbitals for one of the states but for all considered states. The weight, in which each state influences this procedure, can be defined individually. The fact that all states are described by one set of orbitals ensures the orthogonality of the states. This procedure facilitates the description of conical intersections, which will be described in more detail in the following section.

In this thesis, state-averaging is used for all excited state calculations with CASSCF. The influence of both the first excited and the ground state is chosen to be equal.

The active orbitals for CASSCF have to be chosen according to the specific investigated problem. An example for the sensibility of this method and its impact on calculation results will be shown in the following paragraph.

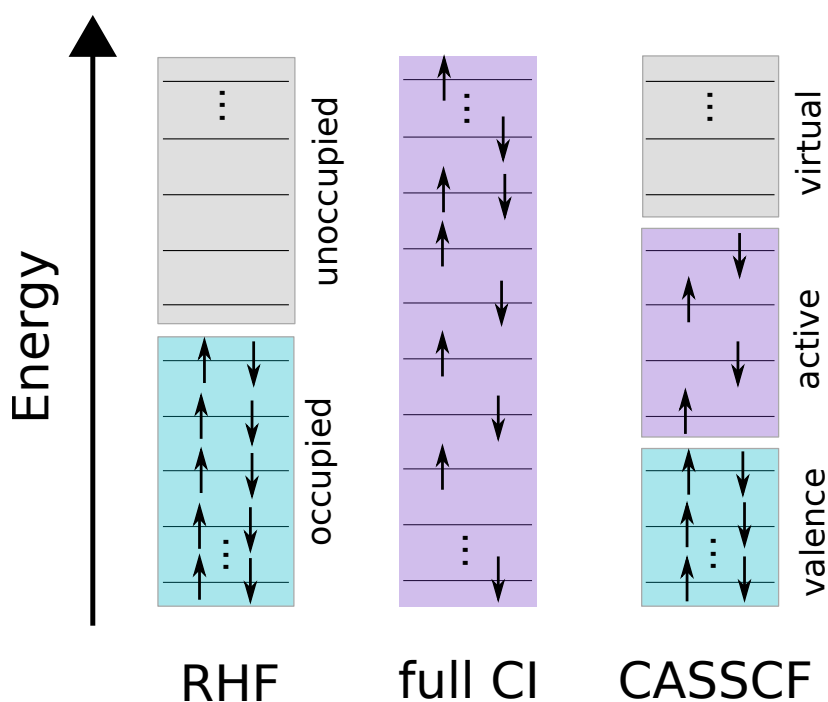


Figure 2.2: Differences between RHF, full CI, and CASSCF. For RHF, each orbital can be either doubly occupied or unoccupied. Full CI has no restrictions regarding the orbital occupancy and all possible excitations are allowed. The combination of both approaches is CASSCF, where the active orbitals are treated as in the case of full CI. Orbitals with lower energy than the active orbitals are always doubly occupied, and orbitals with higher energy are always unoccupied as it would be the case for RHF.

CAS Active Space

The choice of an appropriate active space is essential when applying CASSCF. The best representation is obtained by including all orbitals that take part in the investigated transition. For investigation of deactivation from the first excited state to the ground state in PYP, the active space needs to describe the $\pi \rightarrow \pi^*$ excitation from the ground to the excited state. CAS 12,11, which means CASSCF with 12 electrons and 11 orbitals as active space, fulfils this requirement by incorporating the six highest occupied π -orbitals and the five lowest unoccupied π^* -orbitals.

Long QMMM MD simulations using the QM method CAS 12,11 are computationally very expensive. Reducing the size of the active space is a way to use the available computational resources efficiently. This approach decreases the description of the system. Therefore, it is necessary to control whether the investigated properties of the system are affected.

For this purpose, the size of the active space for the chromophore was reduced step wise from CAS 12,11 to CAS 6,6 with the 6-31G* basis set. For all active spaces, the db isomerisation and sb rotation barriers were calculated in the first excited state using the Gaussian rigid scan method with a basis set of 6-31G*. Convergence was achieved, when the energy difference between two steps was less than 10^{-6} in less than 512 steps. State-averaging was used with equal weight on the ground and the first excited state.

Additionally, permutations of the binding orbitals were performed for CAS 10,9, CAS 8,7, and CAS 6,6 (visualisation of the permuted orbitals can be found in Table 7.6 in the appendix). These were carried out to test the stability of CASSCF active space for the case that a suboptimal orbital was chosen to be part of the active space. An active space was considered stable if a comparable description for the single and the double bond of the chromophore ethylene chain existed. All calculations were performed with the chromophore model pCK⁻, a ketone derivative of the deprotonated *trans-p*-coumaric acid (depicted in Figure 2.3), in vacuum.

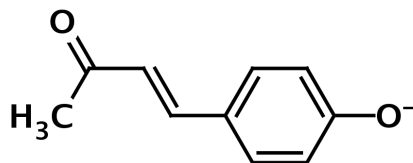


Figure 2.3: Chromophore model pCK[−] which is a ketone derivative of the deprotonated *trans-p*-coumaric acid.

When reducing the size of the active space, the height of the db isomerisation barrier decreased. The height of the sb rotation barrier showed no distinct trend upon reducing the active space size as shown in Table 2.1.

Table 2.1: Barriers for db isomerisation and sb rotation calculated with the rigid scan method implemented in Gaussian.

active space	db barrier (kJ/mol)	sb barrier (kJ/mol)
12-11	24.13	2.625
10-10	17.5	10.8
10-9	16.25	0.604
8-8	10.7	5.4
8-7	11.74	0.866

Permuting orbitals for CAS 10,9 and CAS 8,7 did not produce considerable deviations of the db isomerisation and sb rotation barrier heights. For both active space sizes, the barrier for db isomerisation was notably higher than for sb rotation (Table 7.7 in the Appendix). For CAS 6,6 the results of the permuted active spaces differed significantly. Table 2.2 lists the observed barriers for four permuted CAS 6,6 active spaces.

Table 2.2: Db isomerisation and sb rotation barriers for permuted CAS 6,6 active spaces.

active space	db barrier (kJ/mol)	sb barrier (kJ/mol)
6-6	35.5	—
6-6-a	3.6	5.43
6-6-b	29.12	—
6-6-c	15.46	—

Comparing the bound orbitals for the CAS 6,6 permutations shows that the differences result from an unbalanced description of the chromophore, where one bond, single or double, is described better than the other, for the CAS 6,6 active space permutations 6-6, 6-6-a, and 6-6-c.

Figure 2.4-1 shows bound orbitals of CAS 6-6. Both double and single bond are not described sufficiently, especially the double bond description is weak. Compared with the other two presented permutations the electron density over the double bond is small. This results in a large barrier for db isomerisation of 35.5 kJ/mol whereas sb rotation can occur without crossing any barrier.

In contrast, the description of the double bond for 6-6-a (Figure 2.4-2) is improved as the electron density at the double bond is highest in this permutation compared with both 6-6 and 6-6-c. Thereby the db isomerisation barrier is lowered to 3.6 kJ/mol, which is lower than the corresponding sb rotation barrier with a height of 5.43 kJ/mol.

The orbitals of 6-6-c are evenly distributed over the ethylene chain of the chromophore, which yields a consistent description of single and double bond. The db isomerisation barrier has a height of 15.46 kJ/mol whereas no barrier for the sb rotation is observed.

When reducing the active space of PYP in Gaussian03⁶⁰, it becomes obvious that the algorithms used down to CAS 8,7 are able to exchange orbitals themselves to lower the overall energy of the system. The results, obtained with CASSCF down to this active space, agree with the results for the complete active space qualitatively. The db isomerisation barrier is significantly higher than the sb rotation barrier.

From the reduction down to CAS 6,6 on, the selection of orbitals has to be accurate for the problem at hand. The algorithm changed and does not exchange energetically unfavourable orbitals. Thus an inadequate selection of active space orbitals results in a poor description of the system followed by deficient results. In the example of PYP, the selection of orbitals for CAS 6,6 can result in both a clearly favoured db isomerisation as well as a clearly favoured sb rotation depending on the selected bound orbitals when reducing the active space from CAS 8,7 to CAS 6,6.

Based on this observations, the CASSCF calculations of this thesis were performed using CAS 8,7 and CAS 8,8 for QMMM MD simulations and CAS 12,11 for QM calculations.

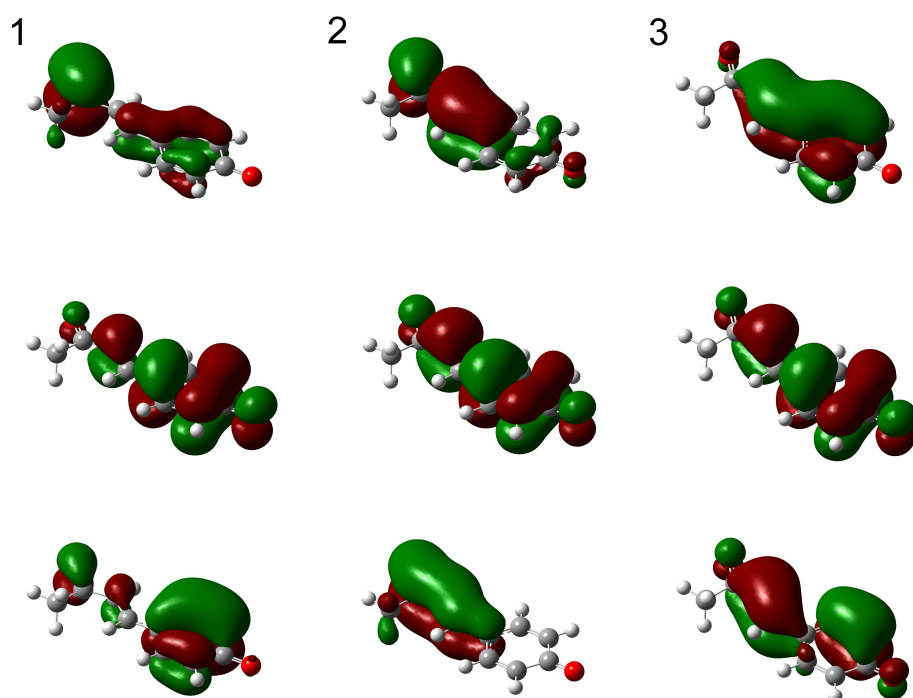


Figure 2.4: The three highest occupied molecular orbitals of the chromophore model pCK⁻ for the CAS 6,6 active space permutations 6-6 (1), 6-6-a (2) and 6-6-c (3). The phenyl ring of the chromophore is always on the lower right side of each image and the ethylene chain on the upper left side.

2.6 Conical Intersection

The QM methods described in this chapter use the Born-Oppenheimer approximation. This approximation uses the different movement speeds of electrons and nuclei as reason to separate the electronic from the nucleic motions. The electrons are assumed to adjust immediately to a changed nuclei conformation. This approach simplifies the Schrödinger equation by only taking into account electronic terms for fixed nuclei positions. This approach is also called adiabatic approximation because the nuclei move on a potential energy surface (PES), defined by the electronic energy for all nuclei configurations, without coupling to other PESs.

The Born-Oppenheimer approximation breaks down when two PESs are energetically close. At this point the non-adiabatic coupling between the PESs can not be ignored. Such a point is called a conical intersection and enables radiationless decay from an excited to a lower energy state in a time scale below one nanosecond, which makes conical intersections important to ultrafast photoinduced processes.

The requirements for conical intersections were described by Teller in 1937⁸⁶ and will be explained in the following passage.

A molecule with N atoms ($N > 2$) has $F = 3N - 6$ degrees of freedom which describe the potential energy surfaces. For a system of x_{nuc} internal coordinates all but two electronic eigenfunctions are defined in the case of degeneracy. The two electronic wave functions of the PESs, ϕ_1 and ϕ_2 , can be chosen in a way that they form a complete orthonormal basis in the remaining subspace with the already known solutions. Close to degeneracy, the eigenvalues of the total Hamiltonian are found within this subspace by examining the Hamiltonian sub-matrix

$$H = \begin{pmatrix} H_{11}(\vec{x}) & H_{12}(\vec{x}) \\ H_{21}(\vec{x}) & H_{22}(\vec{x}) \end{pmatrix}, \quad (2.45)$$

where $H_{12}(\vec{x})$ and $H_{21}(\vec{x})$ contain the non-adiabatic coupling vector \vec{g} and $H_{11}(\vec{x})$ and $H_{22}(\vec{x})$ the gradient difference vectors \vec{h} .

The Eigen values E_1 and E_2 close to the degeneracy are obtained by diagonalising the matrix which leads to

$$E_{1,2} = -\frac{\Delta H}{2} \pm \sqrt{\left(\frac{\Delta H}{2}\right)^2 + H_{21}H_{12}}, \quad (2.46)$$

with $\Delta H = H_{11} - H_{22}$. The square root has to disappear for both states to be degenerate, meaning $E_1 = E_2$. This is fulfilled when

$$\begin{aligned} \Delta H(\vec{x}_{nuc}) &= 0 \\ \text{and} \quad H_{12}(\vec{x}_{nuc}) &= 0. \end{aligned} \quad (2.47)$$

This means, two constraints have to be met: ΔH and $Re(H_{12})$ have to disappear for the real case. The degenerated space for a system with F degrees of freedom is therefore $(F - 2)$ -dimensional.

The number of constraints makes it impossible for diatomic molecules to fulfil these conditions, neglecting the trivial solution of two states that are degenerate for every value of \vec{x} . Instead of a conical intersection the potential energy curves for these systems show avoided crossing, i.e., the potential energy curves never cross.

The term conical intersection arises from the graphical form of the potential energy surfaces. The conditions lead to a two-dimensional space, called the branching space, that is spanned by the gradient difference vector and the non-adiabatic coupling vector. Plotting the energy of the potential energy surfaces over this branching space, in linear approximation the potential energy surfaces form a double cone at the point of degeneracy.

Orthogonal to the two-dimensional branching space a so-called seam space exists, in which the states are degenerate to first order (Figure 2.5). This forms a $(F - 2)$ -dimensional seam for molecules with F degrees of freedom at which each point corresponds to a conical intersection. This means that the system is able to undergo radiationless decay at each point of the seam.

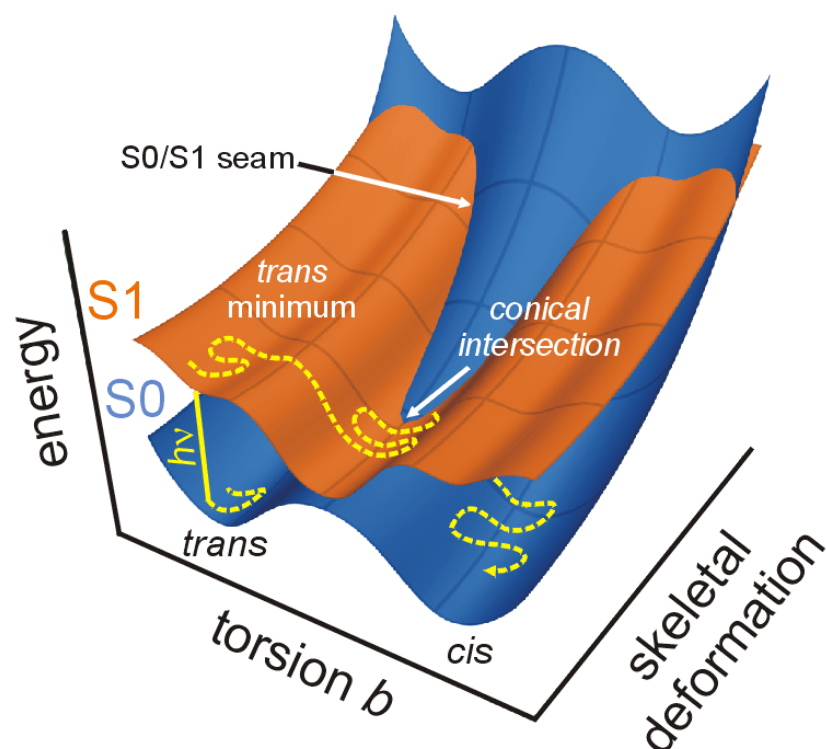


Figure 2.5: The potential energy surfaces of the excited state S1 (red) and the ground state S0 (blue) along the *trans*-to-*cis* isomerisation coordinate b and a skeletal deformation of the bonds. The seam between the two surfaces consists of conical intersections. A typical trajectory is depicted in as yellow line. It starts on the ground state before excitation. On the excited state, it moves from the Franck-Condon-region over a barrier to the db-twisted minimum. From this geometry a hop to the ground state via a conical intersection takes place. On ground state evolution to *trans*- and *cis*-conformation are possible. (Adapted with permission from Groenhof *et al.*²⁵. Copyright 2004 American Chemical Society.)

2.6.1 Surface Hopping

It is important to describe the potential energy surfaces accurately to model the dynamics of a photochemical reaction. As this reaction starts in the excited state and ends in the ground state modelling this crossing consistent with the QM method is necessary.

In this work, Newton’s equations of motion are used to calculate the trajectories of a MD simulation. The quantum-chemical nature of the nuclei is neglected and they are assumed to behave classically. The potential energy surfaces are calculated locally on-the-fly using state averaging for a consistent description of both PESs. Transitions between the PESs at the conical intersection seam are introduced by hopping between the surfaces. This hopping is implemented by using a surface hopping algorithm developed by Groenhof *et al.*⁸⁷ described in the following passage.

A hop from the excited (cipher 1) to the ground state (cipher 2) takes place when the energy difference ΔE between the two states is below a certain threshold and when the CASSCF eigenvector indicates a crossing. These criteria are reassessed at each step i of the simulation. As the CASSCF calculations are carried out using state averaging, the corresponding eigenvectors, C_i^1 for the excited state and C_i^2 for the ground state, are applied to this criterion. The inner product $C_i^2 \cdot C_{i-1}^2$ has to be sufficiently small, whereas the inner product $C_i^1 \cdot C_{i-1}^2$ has to be close to 1 for a crossing of two diabatic surfaces to occur. After these criteria are met, the hop from the excited to the ground state is performed and the rest of the simulation is carried out in the ground state.

This method has the advantage that the diabatic surface is never left, which conserves energy. The method allows surface hops only at the conical intersection which leads to a lower surface hop probability and longer excited state decay times than methods where hops at extended weak coupling regions are allowed. Surface hops in weak coupling regions have been shown to inhibit torsions of the ethylene chain in PYP⁸⁸ and thus suppress the photocycle. Therefore, the above described surface hopping method is applied in this thesis.

Chapter 3

Energy Potential Along the Chromophore - Glu46 Hydrogen Bond

The first structure of PYP was published by Borgstahl *et al.*⁸⁹ in 1995. It had an resolution of 1.4 Å and displayed two amino acids hydrogen bonded to O4a, Tyr42 and Glu46. These hydrogen bonds had a length of 2.71 and 2.69 Å, respectively, which is in the range of normal hydrogen bonds as has been illustrated in section 1.2.

The residues Tyr42 and Glu46 are highly conserved in PYP-like proteins. Kumauchi *et al.*⁹⁰ compared the sequence of 14 PYP structures and found tyrosine at position 42 in 13 structures and glutamic acid at position 46 in 12 structures.

The importance of the hydrogen bonds between the chromophore and Tyr42 and Glu46 has been studied by mutating Tyr42 and Glu46^{19,35,91–93}.

Mutations of both amino acids weaken the hydrogen bonds and slow down the formation of the twisted intermediate on the excited state and the subsequent deactivation to the *cis*-conformer. Brudler *et al.*¹⁹ found a second ground state chromophore conformation for the mutant Y42F by UV/Vis and FT-Raman spectroscopy. The same observation was made by Joshi *et al.*⁹³ from absorption, fluorescence, and FTIR spectroscopy. Both publications state that the chromophore is protonated in the second conformation and Glu46 deprotonated. They constitute the possibility of the protonated chromophore and deprotonated Glu46 with the length of the hydrogen bond connecting the residues. This hydrogen bond has a length of 2.51 Å in the Y42F mutant which is shorter than in the wild type and might offer the possibility for the proton to be bound either to the chromophore or to Glu46. For mutations of Glu46 to glutamine (E46Q) the same qualitative results

as for the wild type were observed in transient-absorption spectra³⁵. The mutation decreases the isomerisation rate and slows down the formation of the *cis*-conformer. The excited state dynamics of the mutant are almost the same as for the wild type. The mutant E46Q was compared to the mutation of Glu46 to alanine (E46A) by Losi *et al.*⁹². They observed significantly slower deactivation from the excited state for E46A. The hydrogen bond between the chromophore and residue 46 is lost for E46A while it is present in E46Q where it has normal hydrogen bond length, 2.87 Å. They conclude that the hydrogen bond to residue 46 mediates the deactivation from the excited state.

The observed lower isomerisation yield of the mutants corresponds to theoretical studies, which presume that the hydrogen bonds to the phenyl oxygen favour deactivations via db isomerisation. This deactivation mechanism is assumed to be the starting event of the productive photocycle of PYP. The missing or weaker hydrogen bonds in the mutants would therefore lead to a lower isomerisation yield. The theoretical concept is described in detail in chapter 4.

Yamaguchi *et al.*⁴⁹ published a structure of PYP in the ground state in 2009 (PDB ID 2ZOH), which was determined by combining X-ray crystallography and neutron diffraction. They were able to achieve a resolution of 1.25 Å (X-ray crystallography) for the heavy atoms and 1.5 Å (neutron diffraction) for the hydrogen and deuterium atoms. 87% of those positions could be resolved. The two hydrogen bonds to the phenyl oxygen of the chromophore are considerably shorter than in the first published X-ray structure. They measure 2.52 Å for the hydrogen bond to Tyr42 (earlier: 2.69 Å) and 2.56 Å for the hydrogen bond to Glu46 (earlier: 2.71 Å), respectively. Additionally, the obtained structure shows two remarkable features of the environment near to the chromophore. That are: In contrast to previous data 1) Arg52 was found to be deprotonated and 2) the position of the proton in the hydrogen bond between the chromophore and Glu46 is in the middle of the bond: Glu46-OE2 - H: 1.21 Å, H - chromophore-O4a: 1.37 Å. The shortness of this hydrogen bond has been reported before but it is the first time that the location of the proton could be resolved. As the observed position is characteristic for an SSHB (see section 1.2) it is interpreted as such by Yamaguchi *et al.*

Saito *et al.*⁹⁴ investigated these results theoretically to gain better understanding of the hydrogen bond between the chromophore and Glu46. They used the structure of Yamaguchi *et al.* and started with the presumption that Arg52 is protonated because of its pK_a which is supported by the

fact that QMMM calculations with deprotonated Arg52 performed by them show a different geometry than X-ray structures. Their study is based on 1) QMMM calculations of the potential energy curves of the hydrogen bonds between the chromophore and Tyr42 and Glu46 and 2) pK_a calculations of the chromophore and Glu46. Their QMMM calculations were carried out using B3LYP/LACVP* as QM and the OPLS2005 force field as MM method. The QM region consisted of chromophore, Tyr42, Glu46 and Thr50 while the remaining part of the protein formed the MM region. They found that the proton was located at a position with 1.0 Å distance to Glu46-OE2. The energy potential showed no further minimum, neither in the middle of the bonding distance nor at the chromophore. The energy potential had the shape of an asymmetric single well in which the proton is located 1.58 Å from the chromophore and 1.00 Å from Glu46. They referred to this bond as LBHB. The pK_a calculations supported this findings as the pK_a of Glu46 is 8.6 while the chromophore has a pK_a of 5.4. These observations suggest a short hydrogen bond with a normal hydrogen position at the donor contradicting the hydrogen position in the middle of the bond observed by Yamaguchi *et al* and their conclusion of this being an SSHB.

The results by Yamaguchi *et al.* are very interesting especially with the contradicting data of Saito *et al.*. To resolve this contradiction, the following chapter will investigate the hydrogen bond between the chromophore and Glu46 by calculating the energy potential along the bond with varying QM region size and Arg52 protonation states.

3.1 Influences on the Hydrogen Bond Energy Potential

To investigate how these two contradicting observations arise we set up different QMMM systems. The QM method chosen was B3LYP/6-31G* with 56 % local DFT and 44 % Hartree-Fock exchange. The MM region was simulated as point charges that were fixed in space with charges taken from the Amber03 force field⁹⁵. With this setup the energy potential along the hydrogen bond between the chromophore and Glu46 was investigated by calculating the energy for each of 25 to 40 positions for the proton evenly spread along the hydrogen bond. Different compositions of the QM region were explored. These calculations are carried out using the unminimised X-ray structure. This structure is assumed to correctly represent the ensemble average positions of the atoms in PYP. The effect of minimising the structure has been

investigated and is shown in section 3.3.

Several factors were found to influence the energy potential and will be discussed separately in more detail in the following paragraphs.

Effects of Electrostatics

The energy potential of the aforementioned hydrogen bond depends on the cut-off that is applied to the point charges surrounding the QM region. The larger the cut-off, the more atoms of PYP are taken into account as point charges and can thus influence the QM region electrostatically. With a cut-off of 1.0 nm the energy potential resembles an asymmetric double well with a distinct minimum for the proton bound to Glu46 and a small minimum at the chromophore. Both minima are separated by a barrier of 20 kJ/mol. By increasing the cut-off the minimum of the energy potential for the proton bound to the chromophore becomes more pronounced which can be perceived as a favoured delocalisation of the proton. (Figure 3.1)

A cut-off of 2.5 nm that considers the whole protein lowers, in case of neutral Arg52, the barrier for proton transfer from Glu46 to the chromophore by about 10 kJ/mol and originates a local minimum for the proton bound to the chromophore. This minimum lies 5 kJ/mol higher in energy than the minimum at Glu46. For protonated Arg52 the barrier is lowered by 10 kJ/mol and no minimum at the chromophore is formed.

To further investigate the effects of the electrostatic interactions of the MM part onto the QM region the ONIOM QMMM scheme in Gromacs was used. Thus, the QM region resembled that of the normal QMMM calculations but the charges of the surrounding protein did not influence the QM region atoms. For single point calculations this setup is the same as a calculation of the cluster model of the QM region.

The resulting potential energy curve for neutral Arg52 shows no double well potential but a single minimum for the hydrogen at Glu46. The Arg52 protonation state still influences the energy potential in about the same scale for both ONIOM and normal QMMM but the deprotonation does not give rise to a double well potential as depicted in Figure 3.2.

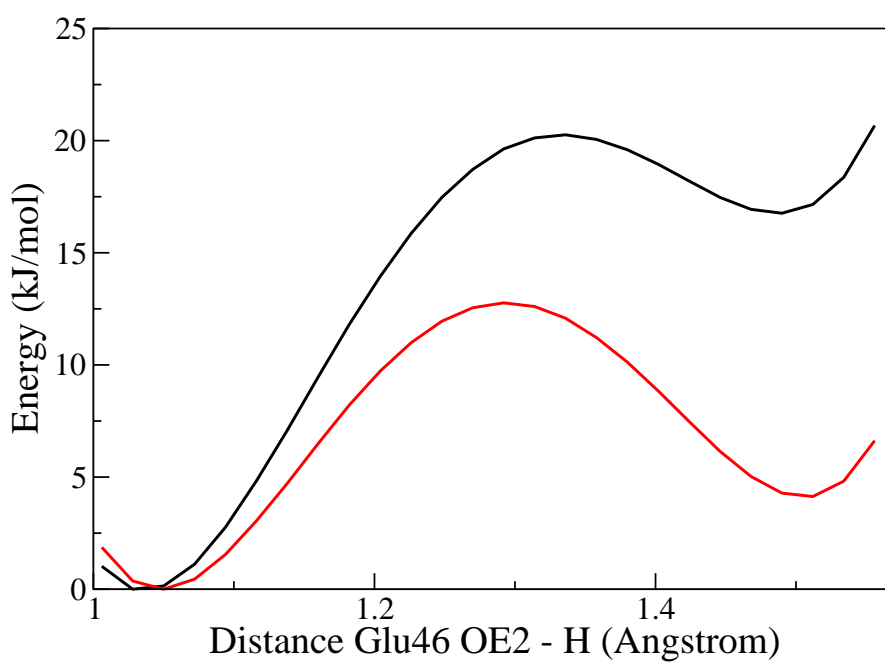


Figure 3.1: Comparison of the energy potential with 1.0 nm (black) and 2.5 nm cut-off (red) for structure A with QM region 3 (see table 3.1) and neutral Arg52. A larger cut-off which considers the electrostatic effects of the whole protein unto the QM region changes the shape of the energy potential significantly. Not taking into account parts of the protein results in too large barrier heights.

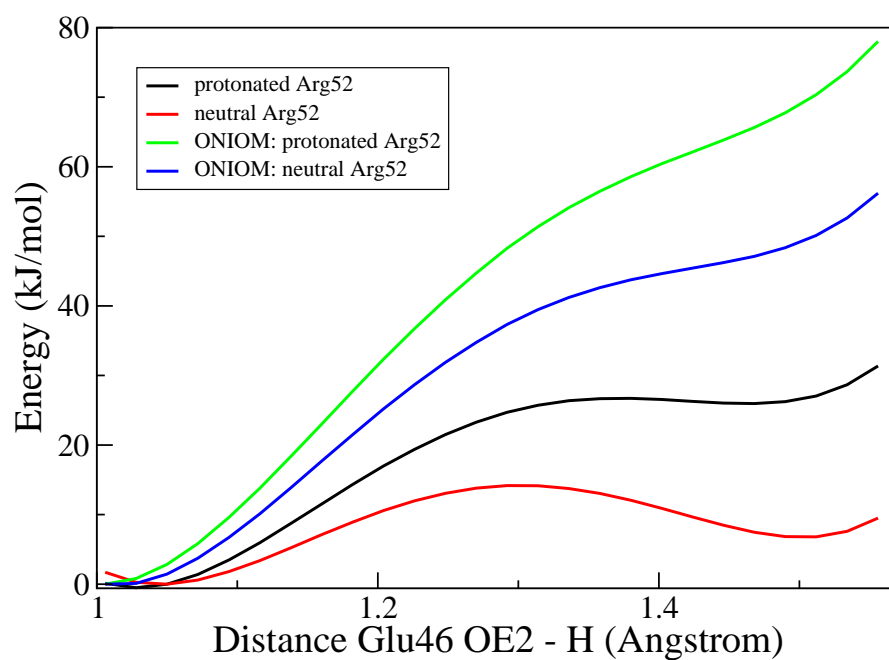


Figure 3.2: Comparison of the energy potential of calculations using ONIOM and the normal QMMM scheme for minimised structure A with QM region 3 and protonated/unprotonated Arg52.

Effect of Protein Conformation

The crystal structure published by Yamaguchi *et al.* contains two structures A and B that differ in the side chain position of seven amino acids (Ser8, Gln41, Ile58, Phe79, Ser117, Tyr118 und Lys123). All these amino acids are located at the protein surface except for Gln41 which is situated in the binding pocket. As Saito *et al.* did not specify which of the structures in the RCSB entry 2ZOH they chose, the energy potentials of both structures were calculated and compared. Structure B favours the delocalisation of the proton more than structure A. (Figure 3.3)

By calculating the energy potential along the hydrogen bond between Glu46 and the chromophore for all variations of structures A and B the cause for the difference was identified. The different side chain orientation of Lys123 is responsible for the difference of the energy potential of structures A and B. Lys123 is located about 14 Å away from the chromophore on the surface of the protein. In structure A the positively charged Lys123 side chain points into the bulk water while it point towards the protein in structure B and forms a hydrogen bond to Ala27. The difference in orientation for the remaining six amino acids shows no effect on the energy potential. The influence of Lys123 on the QM region is due to the fact that of all residues existing in multiple conformations only Lys123 is charged. (Figure 3.4)

To cross validate this result the energy potentials were also calculated for structure B with the coordinates from each individual differing amino acid from structure A. These calculations confirmed the observations from the first calculation setup by identifying Lys123 to be the origin of the differences in the hydrogen bond energy potential.

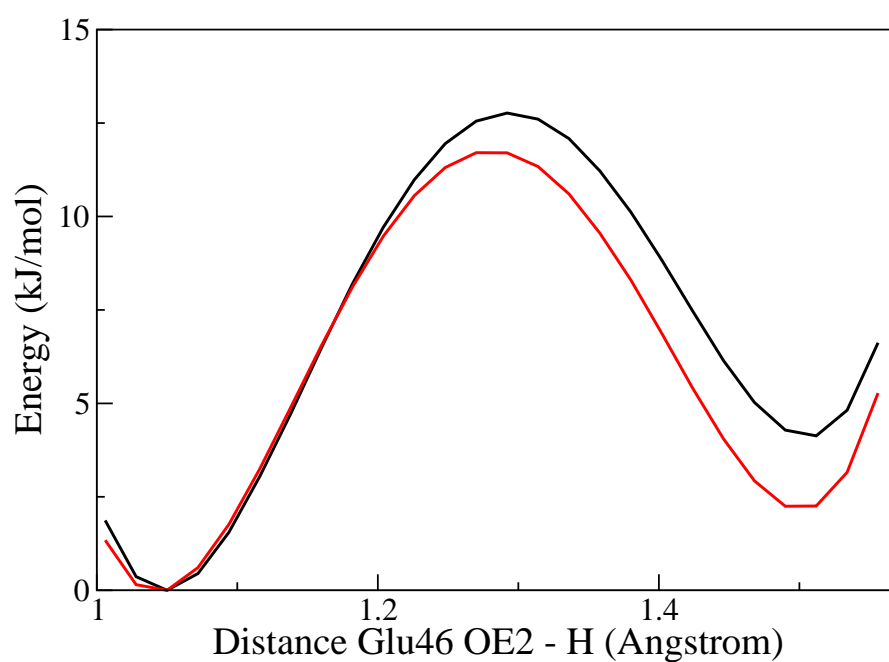


Figure 3.3: The energy potentials of structures A (black) and B (red) for QM region 3 with neutral Arg52. The differing orientations of seven amino acids side chains in structure B compared to structure A cause a preference of the delocalised proton.

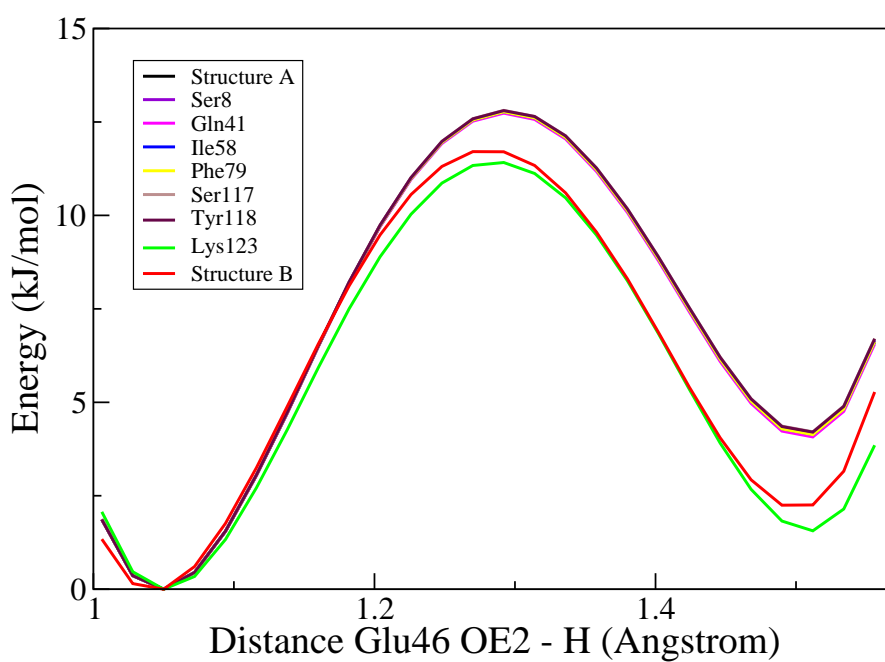


Figure 3.4: Comparison of the energy potential of structures A, B and single differing amino acids for QM region 3 with neutral Arg52. Structure A and B differ in the positions of seven amino acid side chains. When the different coordinates from structure B for one of these amino acid side chains is used in structure A only for Lys123 the different potential energy could be observed.

Composition of QM Region

To investigate how the incorporation of selected amino acids into the QM region affects the energy potential six different compositions of the QM region were analysed for each of the two protonation states of Arg52. The QM region consistently contained the chromophore, Tyr42 and Glu46. Cys69, Thr50 and Arg52, either protonated or neutral, were added to this basis system in different combinations (QM regions 1 - 6, see table 3.1).

Table 3.1: Numbering and selected amino acids of the different QM regions for which the potential energy curve was calculated to compare the effect of choosing different compositions of the QM region

QM region	amino acids
1	chromophore, Tyr42, Glu46, Arg52
2	chromophore, Tyr42, Glu46, Thr50
3	chromophore, Tyr42, Glu46, Arg52, Cys69
4	chromophore, Tyr42, Glu46, Cys69
5	chromophore, Tyr42, Glu46, Thr50, Cys69
6	chromophore, Tyr42, Glu46, Thr50, Arg52 Cys69

The energy potentials obtained in these calculations have either the shape of an asymmetrical single well with only one minimum for the proton bonded to Glu46 or show a minimum for each of the two possible proton positions at the chromophore or at Glu46. For the latter energy potentials the minimum at Glu46 is more pronounced than the minimum at the chromophore.

The most prominent effect on the hydrogen bond energy potential was caused by the Arg52 protonation state independent of the QM region composition.

For protonated Arg52 the energy potential had the shape of an asymmetrical single well. The minimum was at Glu46 and no or only a small minimum with a depth of less than 5 kJ/mol was observed at the chromophore. The barrier was 23–33 kJ/mol high depending on the QM region composition. The positive charge of the protonated Arg52 near the chromophore stabilises its negative charge, which therefore needs less stabilisation by the proton from the hydrogen bond to Glu46. The result is that the proton is covalently bound to Glu46.

After deprotonation of Arg52 the picture changed. The energy potential was

shaped as asymmetrical double well with a low barrier separating the two minima. All QM regions, except for QM region 2, showed a distinct minimum at the chromophore which was 4–10 kJ/mol higher in energy than the minimum at Glu46. The barrier separating both minima was 12–16 kJ/mol high. The energy potential for QM region 2 was considerably higher in energy than for the other five QM region compositions. The barrier height was 22 kJ/mol and the minimum at the chromophore 19 kJ/mol above the minimum at Glu46. Neutral Arg52 does not stabilise the negative charge of the chromophore. The negative charge is localised on the phenyl oxygen of the chromophore which attracts the proton. This attraction results in the appearance of the hydrogen bond energy potential with two distinct minima. The barrier between both minima is low enough to be crossed by the proton which might result in a delocalisation. Section 3.2 will address the probability distribution of the proton position.

Including the hydrogen bond between the chromophore and the Cys69 backbone into the QM region leads to an asymmetrical double well with two distinct minima. This could be caused by a too strong MM description of this hydrogen bond which is to be considered in future work. A strong hydrogen bond to the carbonyl oxygen would lead to better stabilisation of the negative charge on the chromophore's ethylene chain compared to a weaker hydrogen bond in the same position. This might lead to a charge shift from the phenyl ring to the ethylene chain which would result in less negative charge on the phenyl oxygen. As a result the attraction of the phenyl oxygen for the proton is reduced. By including the mentioned hydrogen bond to Cys69 into the QM region the hydrogen bond strength would be reduced and the negative charge of the phenyl oxygen increased. As a result the shape of the energy potential becomes an asymmetrical double well potential.

Thr50 forms a hydrogen bond to Tyr42 and thus strengthens the hydrogen bond between Tyr42 and the chromophore. This reduces the negative partial charge on the phenyl oxygen. As a result the proton is attracted stronger to Glu46 and the energy potential shows only one minimum at Glu46. Thr50 has a strong impact on the energy potential. Especially when neither Cys69 nor neutral Arg52 are part of the QM region. In this case the reduction of the negative partial charge on the chromophore has stronger impact on the energy potential. (Figure 3.5)

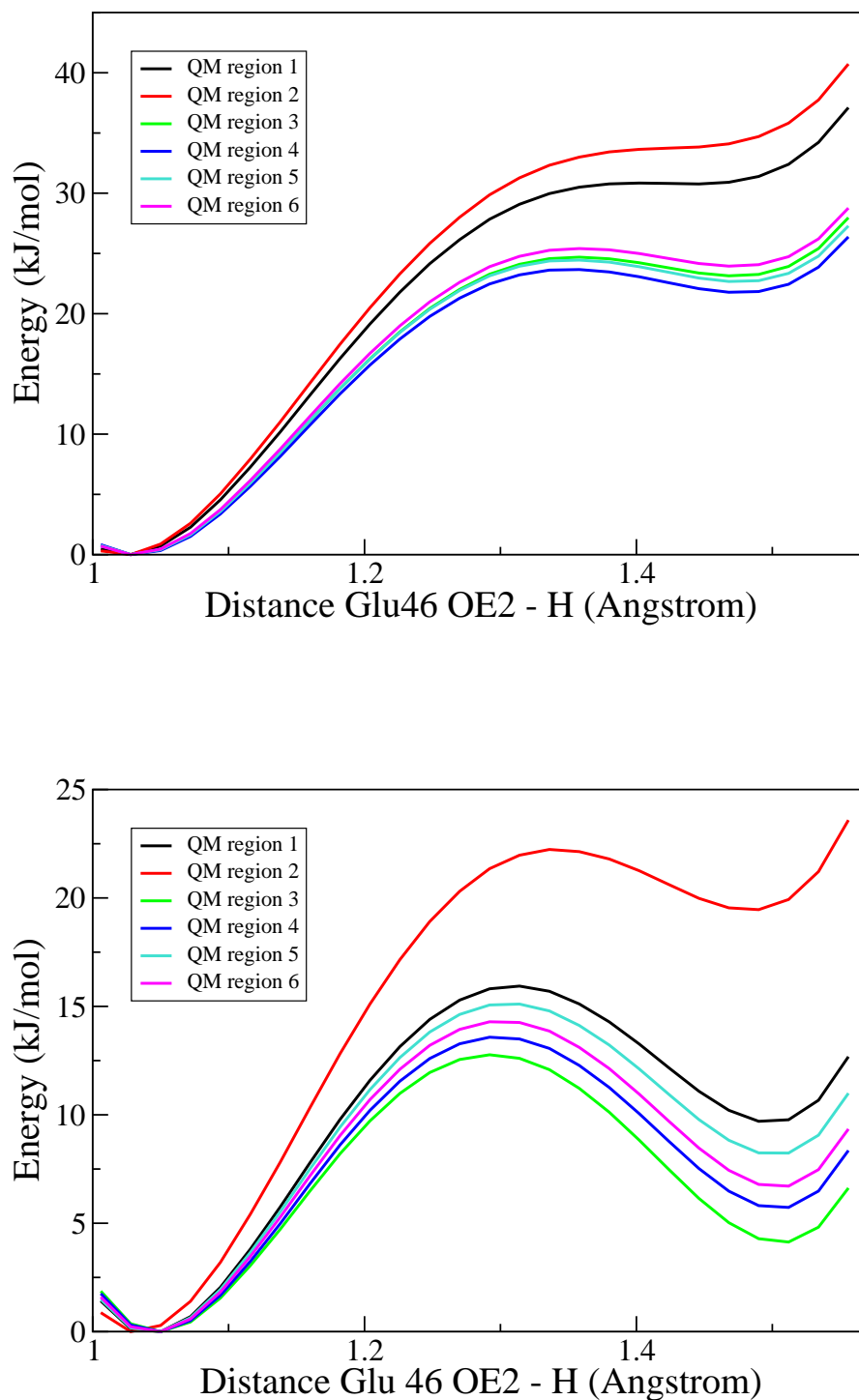


Figure 3.5: Energy potentials along the hydrogen bond of all six QM regions specified in Table 3.1 for protonated (up) and neutral (down) Arg52. Structure A and a cut off radius of 2.5 nm were used to calculate the potentials.

3.2 Location of Hydrogen Atom

The shape of the wave function of the proton along the hydrogen bond between O4a and Glu46 depends on the shape of the potential energy curve for the proton position. The wave function was calculated by solving the time-independent Schrödinger equation with a program provided by Ludger Inhester¹.

The wave functions were calculated for the potential energy curves of QM region 3 in structure A with a cut-off of 2.5 nm for both protonated and neutral Arg52.

The wave function for protonated Arg52 has one maximum at the position of the proton bound to Glu46. In contrast, the computed wave function for neutral Arg52 shows a maximum at Glu46 with a broad shoulder at the position of the proton bound to the chromophore.

From the wave functions the probability density for the proton position along the hydrogen bond was calculated. For protonated Arg52 the probability to find the hydrogen bonded to the chromophore is almost zero. The expectation value for the proton position in this energy potential is 1.08 Å from Glu46 and 1.5 Å from the chromophore. Notably, the proton has a probability to be bound to the chromophore in the potential energy curve calculated for neutral Arg52. In this case, the expectation value is 1.2 Å from Glu46 and 1.38 Å from the chromophore. These values agree well with the experimentally measured distances between the proton and Glu46 and the chromophore of 1.21 Å and 1.37 Å, respectively, by Yamaguchi *et al.* (Figure 3.6)

The vibrational energy for the ground states of the two considered potential energy curves are 13.23 and 9.63 kJ/mol for protonated and neutral Arg52, respectively. The flatter potential energy curve for neutral Arg52 therefore provides an energy gain of 3.6 kJ/mol.

¹MPI BPC Göttingen

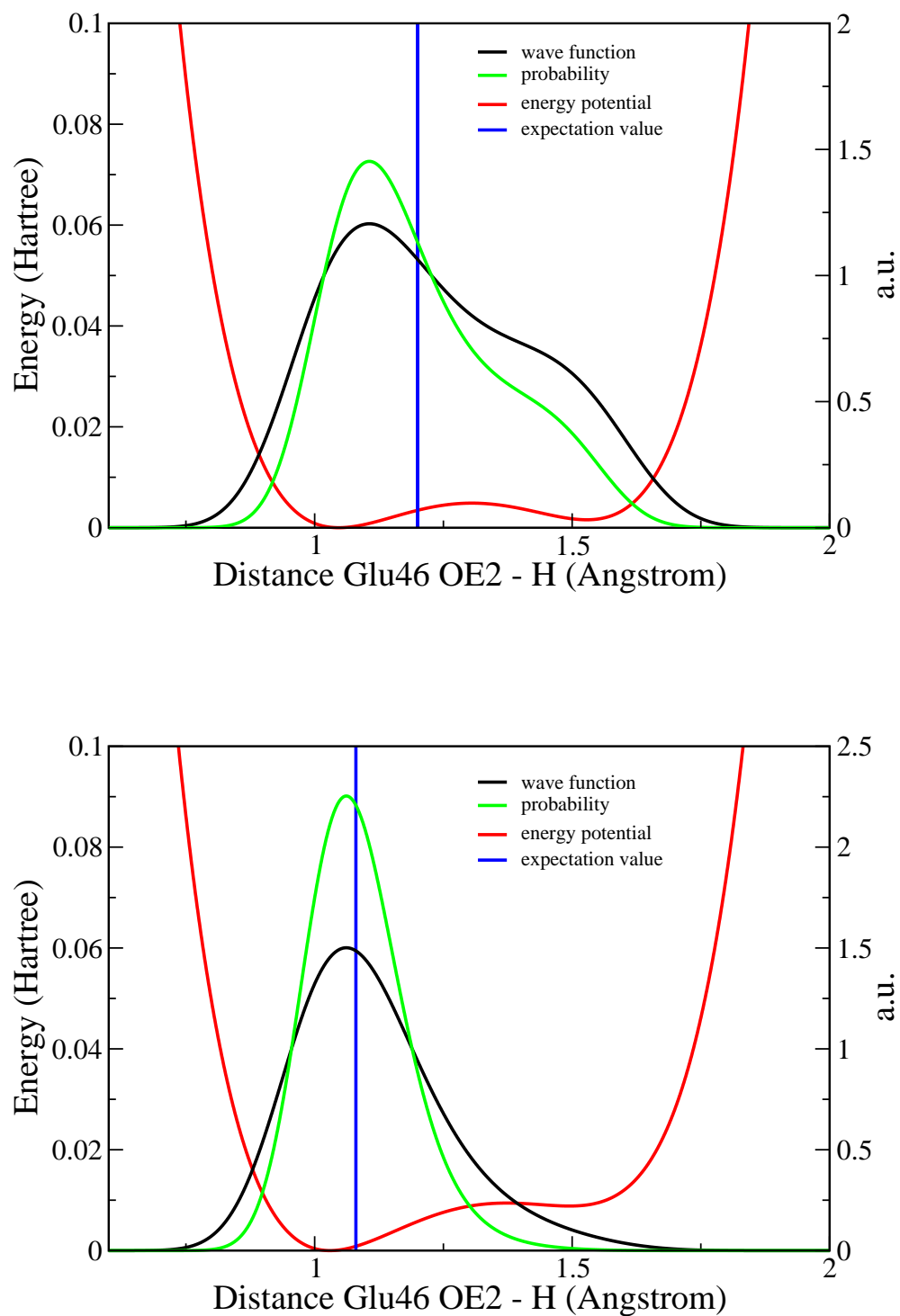


Figure 3.6: The energy potentials (red), wave functions (black), probability densities (green) and expectation values (blue) along the hydrogen bond between the chromophore and Glu46 for neutral (up) and protonated Arg52 (down).

3.3 Minimised Structures

All previous calculations were done using the non-minimised crystal structure 2ZOI allowing to compare the obtained results to the findings of Yamaguchi *et al.* based on the assumption that this structure represents the ensemble average of the atom positions in PYP.

Saito *et al.* minimised the whole protein before calculating the energy potential. Their QM region was additionally minimised for each structure with altered proton position. To investigate the influence of the structure relaxation, the potential energy curves were calculated with the system minimised with steepest descent to machine precision with QM region 3 for both neutral and protonated Arg52 and for both structure A and structure B. The QM method used was again B3LYP/6-31G* with adjusted HF and DFT exchange. With these structures the same calculations as described in section 3.1 were performed.

The energy potential for the system with neutral Arg52 and the side chain positions of structure A before the minimisation shows an asymmetrical double well shape as the unminimised structure but not as pronounced. Both results differ in the barrier height and the height of the minimum at the chromophore, which are about 2 and 4 kJ/mol higher for the minimised structure. (Figure 3.7)

The same effect can be observed in calculations with protonated Arg52 with a minimised structure. Barrier and 'minimum' are higher than in the calculations with unminimised structures but the qualitative picture remains the same.

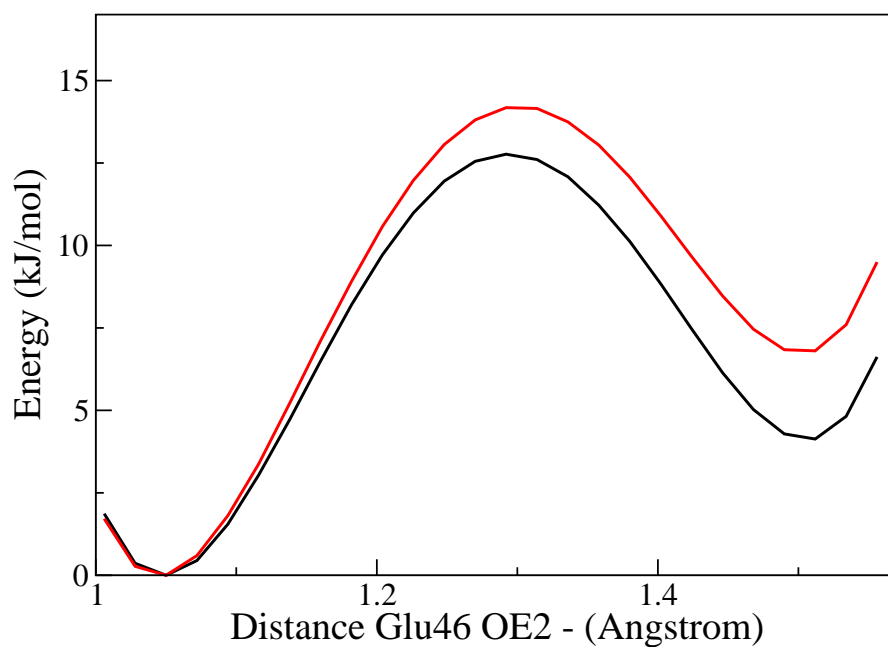


Figure 3.7: Comparison of the energy potential of minimised (red) and unminimised (black) structure A for QM region 3 with neutral Arg52. Minimising the system results in the same qualitative picture with a larger barrier height and less pronounced minimum at the chromophore.

3.4 Discussion

The aim of this chapter was to ascertain the possibility of reproducing the experimentally observed SSHB between the chromophore and Glu46 with QMMM calculations and to identify the conditions under which the SSHB is formed. For neutral Arg52 the expectation value of the proton position along the hydrogen bond is in agreement with experimental measurements. The shape of the energy potential curve was found to be highly influenced by environmental parameters and the model assumption underlying the quantum chemical calculations. A cautious investigation of these dependencies revealed an appropriate choice of QM region.

The observation of neutral Arg52 by Yamaguchi *et al.* contradicts current assumptions that Arg52 in PYP and arginine residues in proteins in general are protonated^{96,97}. The second observation is the position of the proton in the middle of the hydrogen bond between the chromophore and Glu46. This observation can be based on two events. The first possibility is that the measured position is the actual position of the hydrogen in all PYP proteins in the crystal. The second possibility is that of an average of an even distribution between the two proton positions, bound to either Glu46 or the chromophore.

By calculating the pK_a of Arg52 and performing QMMM calculations with protonated and deprotonated Arg52 Saito *et al.* concluded that Arg52 is protonated. Thus, they used protonated Arg52 in all their calculations. The energy potential along the hydrogen bond between the chromophore and Glu46 was calculated for a QM region containing the chromophore Tyr42, Glu46 and Thr50. It is an asymmetrical double well and shows a single minimum for the hydrogen bonded to the chromophore.

To clarify the different observations of both groups both protonation states of Arg52 were used in this work to compare their effects.

The results of this work agree with Yamaguchi and Saito *et al.* depending on the chosen protonation state of Arg52.

For calculations with protonated Arg52 the energy potentials show one minimum for the proton bonded to Glu46. The expectation value for the proton position is 1.08 Å from Glu46 and the minimum of the energy potential is found 1.028 Å from Glu46. This observation agrees with Saito *et al.* who found the minimum at 1.00 Å from Glu46 in his calculations with protonated Arg52.

Neutral Arg52 introduces a second minimum for the hydrogen at the chromophore and lowers the barrier separating the two minima. This leads to

a shift of the expectation value towards the chromophore compared to the calculations with protonated Arg52 to a position 1.20 Å from Glu46 and 1.38 Å from the chromophore. These values agree with the observations of Yamaguchi *et al.* for the case that their measurements show the average position of the hydrogen in all PYP molecules of the crystal instead of its actual position. Whether and how the delocalisation of the proton in the hydrogen bond between the chromophore and Glu46 effects the structure of PYP is an interesting subject for further research.

The effect of differing size and composition of the QM region can not be disregarded. Isborn *et al.*⁹⁸ showed in a recent study that absorption spectra of the PYP chromophore in vacuum, water and the protein calculated from QMMM MD simulations do not converge until large QM regions are used. They found that all calculated absorption bands were blue-shifted compared to experimental spectra.

Their MM part was described with the Amber force field and B3LYP with the 6-31G basis set was used for the QM region. They subsequently enlarged the QM region and convergence was observed for including 40 water atoms for the chromophore in water. For the chromophore in PYP the largest QM region they studied contained 723 atoms. This QM region shows a significant red-shift towards the experimental value compared to the other investigated QM regions and this result was confirmed by comparing the absorption energies from three snapshots of the MD with absorption energies calculated for the same snapshots with describing the whole protein and its counterions with QM. From these observations they concluded that only very large QM regions are able to account for electrostatic interactions and the polarisation effects that are needed to adequately describe the photon absorption of the chromophore.

The high sensitivity of the calculated absorption spectra of Isborn *et al.* regarding the QM region size, poses the question whether a QM region of similar size is needed for the here investigated hydrogen bond potential energy curves. It is presumable that absorption spectra are more sensitive towards the QM region size than potential energy curves. Thus, the here presented qualitative results are not influenced. It would be interesting to confirm this assessment by further calculations with a larger QM region.

Our results regarding the QM region composition are in analogy to the findings of Isborn *et al.* whose study showed that the inclusion of the hydrogen bond between the Cys69 N and chromophore O1 into the QM region induces a red-shift in the absorption spectra which was not observed for calculations without this hydrogen bond. Also in the calculations performed

here, a qualitative different hydrogen bond potential energy curve is found. These two findings can be explained by the following mechanism: Positively charged MM atoms near the QM region, like the proton bound to the Cys69 N, lead to a overpolarisation of the QM electron density⁹⁹⁻¹⁰¹. This overpolarisation results in a changed potential energy curve for the hydrogen bond as well as a red-shifted photoabsorption spectrum.

In summary the findings of Yamaguchi and Saito *et al.* are in agreement with the here presented calculations. In particular, the presumption of the Arg52 protonation state and the SSHB interpretation by Yamaguchi is supported.

The reason for the Arg52 deprotonation is not known. The following mechanisms can not explain the deprotonation alone.

A possible explanation for the deprotonation of Arg52 is the stabilisation by the delocalisation of the proton in the hydrogen bond between O4a and Glu46. The energy gain computed in this work for going from the localised proton (protonated Arg52) to the delocalised proton (neutral Arg52) is 3.6 kJ/mol. The pK_a of Arg52 would have to be lowered by 3, which would correspond to about 18 kJ/mol at pH 9 used in experiment.

Further, the crystallisation can be the reason for the Arg52 deprotonation. The environment of Arg52 in the crystal contains seven water molecules at the border between two PYP molecules. This environment is hydrophobic in contrast to the hydrophilic environment in case of the soluted protein. Theoretical calculations of the pK_a changes from soluted to crystallised PYP by Boggio-Pasqua (personal communications) yielded that the pK_a of Arg52 is lowered by half a unit.

From our calculations neither the delocalisation of the proton nor the crystallisation of the protein or a combination of both can account for the energy needed to stabilise the deprotonated Arg52. Further simulations with included delocalisation of the proton might give insights to possible destabilisation effects leading to Arg52 deprotonation.

Further effects of the protonation state of Arg52 will be discussed in the following chapter.

Chapter 4

Influence of Arg52 Protonation State on PYP Dynamics and Chromophore Isomerisation

The role of Arg52 in PYP was subject of several previous experimental and theoretical studies.

X-ray structures of the wild type and the R52Q mutant, where Arg52 is mutated to the neutral amino acid glutamine, were compared by Shimizu *et al.*²¹ with focus on the chromophore binding site. For the latter they observed a cavity near the chromophore, which was occupied by two water molecules. These water molecules were part of a hydrogen bond network which connected Arg52 to Tyr98 replacing the hydrogen bond existing between these residues in the wild type. The hydrogen bonds between the chromophore and Tyr42 and Glu46 were found unchanged.

Excited state QMMM MD simulations of the R52Q mutant were carried out by Groenhof *et al.*⁸⁷ They studied the deactivation of the chromophore from the excited state. The deactivation took place predominantly by single bond rotation, which is in contrast to previous simulation with the wild type protein by the same researchers²⁵ where only double bond isomerisation as deactivation pathway was observed. The photochemical processes were slower in the mutant than in the wild type and the quantum yield was lower, 0.2 for R52Q compared to 0.3 for the wildtype.

Their quantum yield and kinetics match with experimental studies. Changanet-Barret *et al.*^{34,35} used transient absorption spectroscopy on the wild type and R52Q. They observed a similar relaxation pathway for both proteins. The excited state lifetime of the mutant is increased by the factor 3 whereas the following decay from the excited state is slowed down. The observed quantum yield for was 0.19 for the mutant and 0.31 for the wild type. The

observations were assigned to changed protein conformation and, resulting from this conformational change, to fewer reactive intermediates.

Sindhikara *et al.*¹⁰² theoretically investigated the cavity near the chromophore in R52Q. They calculated the probability for a hydronium ion, which might replace the positive charge of the mutated arginine, in this cavity. Compared to the bulk this probability is high but though the low concentration of hydronium ions at physiological pH the occupation of the cavity by a hydronium ion is relatively low.

To investigate which Arg52 protonation state agrees best with experimental data besides the structure published by Yamaguchi *et al.*, MD simulations with both protonation states were performed for the present investigation. The calculations were carried out for PYP in a water box with physiological salt concentration (0.150 mol/liter)¹⁰³. The relevant atoms of the chromophore and the surrounding amino acids are depicted in Figure 4.1.

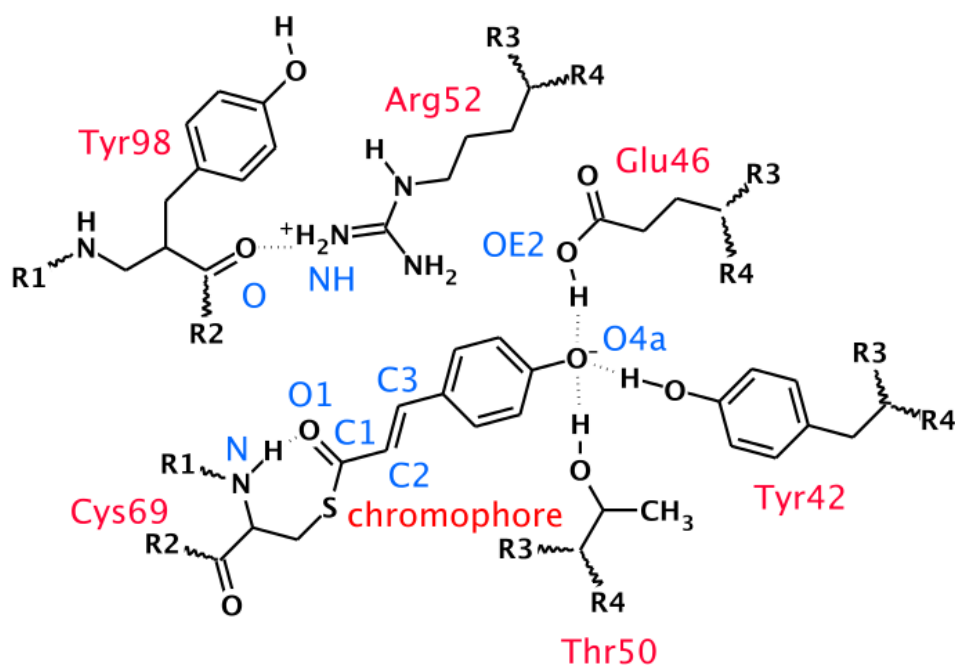


Figure 4.1: Illustration of the amino acids forming the chromophore binding pocket that are relevant in this work with identification of the atoms that are specifically mentioned in the descriptions of this chapter.

4.1 Dynamics

As starting structure PDB entry 2ZOI was used. This model was obtained by Yamaguchi *et al.*⁴⁹ based on both X-ray and neutron diffraction data. Missing hydrogen atoms were added to the protein structure using the Gromacs tool `pdb2gmx`. The protein molecule and 6280 TIP3P water molecules, 24 Na⁺ and 19 Cl⁻ ions for PYP with protonated Arg52 and 6279 TIP3P water molecules, 26 Na⁺ and 19 Cl⁻ ions for PYP with neutral Arg52 were contained in a periodic box and described with the Amber03 force field. In case of protonated Arg52 the C-terminus was accidentally capped with a neutral end group resulting in one additional Na⁺ ions for the system with neutral Arg52 which has a negative end group. The atom charges for the chromophore and neutral arginine were calculated using MP2 and the RESP fitting protocol. Equilibrium bond length and angles were obtained from MP2 optimised chromophore conformation with standard force constants used in the AMBER03 force field. Dihedrals were taken from the AMBER03 force field. All parameters used for the chromophore can be found in the appendix in Tables 7.1 to 7.3 with Figure 7.1 depicting the atom names. The system was minimised using the steepest descent integrator of GROMACS 4.0.5 with a maximum step size of 0.01 nm and converged to machine precision within 132 steps for PYP with protonated Arg52 and 3652 steps with neutral Arg52. F_{max} was 3879,19 and 79.52 kJ mol⁻¹ nm⁻¹, respectively. Subsequently, six MM MD simulations of 100 ns with a step size of 0.002 ps were performed for each protonation state. A different initial velocity distribution was initiated for each simulation by using different values for `gen_seed` in Gromacs. During the MD simulation the PYP molecule was fitted to the reference structure, the output of the preceding minimisation, at each step. The cut-off for van der-Waals interactions was 1.0 nm and for the protein and the solvent the Berendsen thermostat was used independently to keep the system at 300 K.

Protonated Arg52

The minimised structure showed an RMSD of 0.09 nm to the crystal structure for the key residues Tyr42, Glu46, Thr50, Arg52, Cys69, Tyr98 and the chromophore and 0.13 nm for the backbone of the whole protein. The hydrogen bonds between the chromophore O4a and Tyr42 and Glu46 both measured 2.6 Å in the minimised structure which is a slight elongation compared to the crystal structure where these hydrogen bonds measure 2.52 and 2.56 Å, respectively. The structure of the chromophore binding pocket is conserved by the minimisation (Figure 4.2-left).

All six MM MD simulations of PYP with protonated Arg52 in water show

similar behaviour. This will be described on the basis of one of these simulations.

Throughout the MM MD simulation the conformation of the chromophore binding pocket stayed intact as well. A comparison of a snapshot from the MM MD with the minimised starting structure is depicted in Figure 4.2-right.

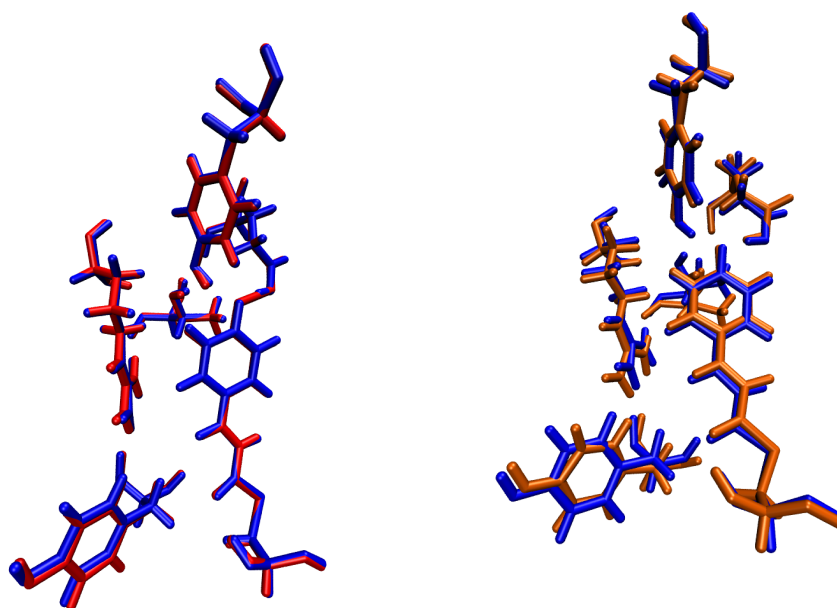


Figure 4.2: Comparison of the minimised structure with protonated Arg52 (blue) with the X-ray structure 2ZOI (red) on the left side. The right picture depicts the minimised structure (blue) and a snapshot from the MM MD simulation (orange) both with protonated Arg52. The chromophore binding pocket is conserved by the minimisation and the MM MD simulation.

The protein structure of PYP with protonated Arg52 was stable during the simulation. The RMSD with respect to the minimised structure was below 0.2 nm in case of the backbone. When taking into account the whole protein except for the chromophore, the RMSD was below 0.3 nm for most of the simulation time as depicted in Figure 4.3.

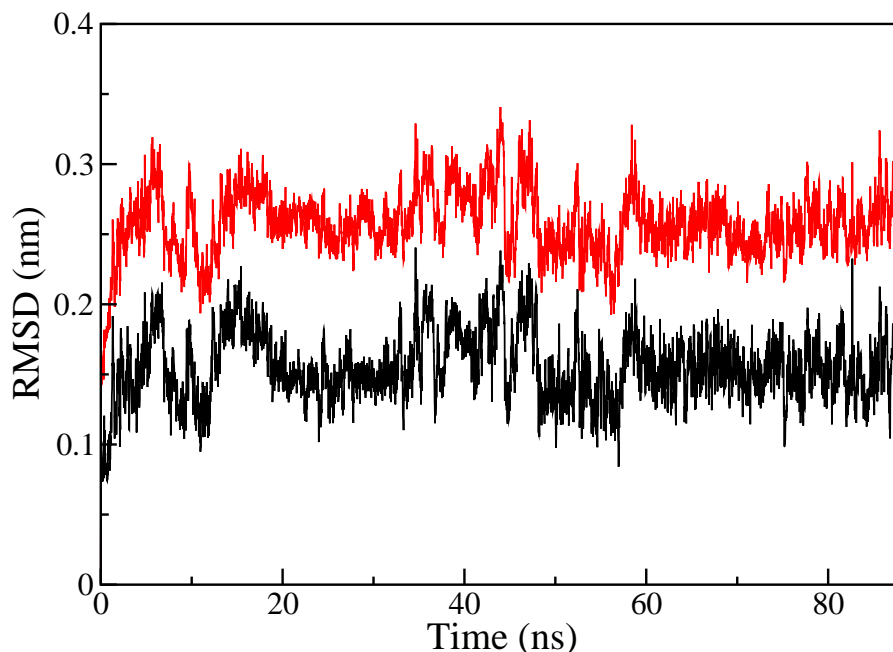


Figure 4.3: Depiction of the RMSD for the backbone (black) and the whole protein except the chromophore (red) throughout the MM MD simulation with protonated Arg52. The minimised structure was used as the reference structure.

The length of the hydrogen bonds between the chromophore O4a and Tyr42 and Glu46 changed during the MM MD simulation. These adopted a length of about 2.8 and 2.7 Å, respectively (Figure 4.4). This is due to the parametrisation of the Amber03 force field which does not contain short hydrogen bonds. Therefore, MD simulations with chromophore, Tyr42 and Glu46 in the QM region were carried out as described in Section 4.1.1.

In contrast to published X-ray structures, Thr50 formed a hydrogen bond to the chromophore O4a for most of the simulation time. A hydrogen bond to Tyr42 as observed experimentally in crystals^{18,22,49} could not be observed in the theoretical simulations carried out for PYP in solution presented in this work.

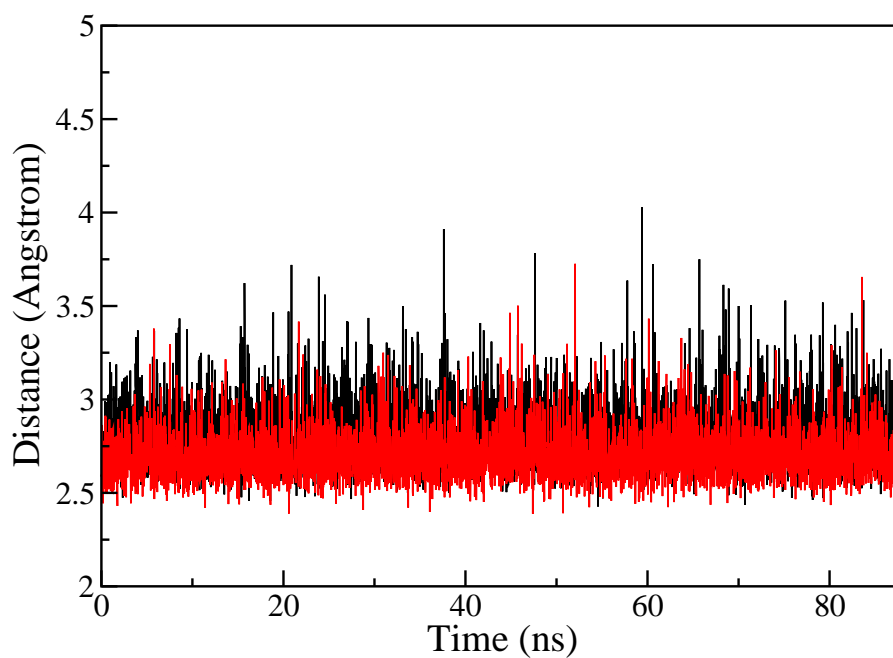


Figure 4.4: During the MM MD simulation with protonated Arg52 the length of the hydrogen bonds between the chromophore O4a and Tyr42 (black) and Glu46 (red) evolved around 2.8 and 2.7 Å, respectively.

Neutral Arg52

The six MM MD simulations of PYP with deprotonated Arg52 in water shared one remarkable feature. The side chains of Arg52 and Tyr98 moved away from their respective positions in the starting structure. In five MM MD simulations a water molecule entered the chromophore binding pocket and bridged the Thr50 side chain to O4a. One of the simulations showed a destabilisation of the chromophore binding pocket with movement of the chromophore. All this features are described in the following using the MM MD simulation that comprises all three observations.

After minimisation, the initial structure for the MM MD simulation with neutral Arg52 was consistent with the crystal structure with an RMSD of 0.33 nm for the protein backbone. The hydrogen bonded network surrounding the chromophore was intact and the amino acids forming the chromophore binding pocket were found at positions in agreement with the ones observed in the crystal structure. The conformation of the chromophore binding pocket is depicted in Figure 4.5.

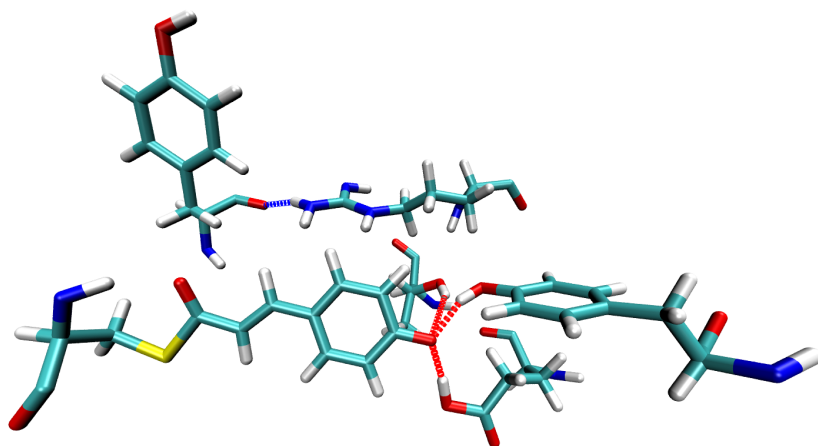


Figure 4.5: Positions of the chromophore, Tyr42, Glu46, Thr50, Arg52, Cys69 and Tyr98 in the initial structure of the MM MD simulation. The positions correspond to the positions observed in the crystal structure 2ZOH.

During the simulation with neutral Arg52, its side chain and the side chain of Tyr98 moved away from their initial positions. The hydrogen bond between Arg52 NH2 and Tyr48 O was broken and both side chains relocated to positions further apart than shown in X-ray structures (Figure 4.6). The side chains of Arg52 and Tyr98 formed transient hydrogen bonds to other parts of the protein which were not stable, resulting in considerable movement of these two residues.

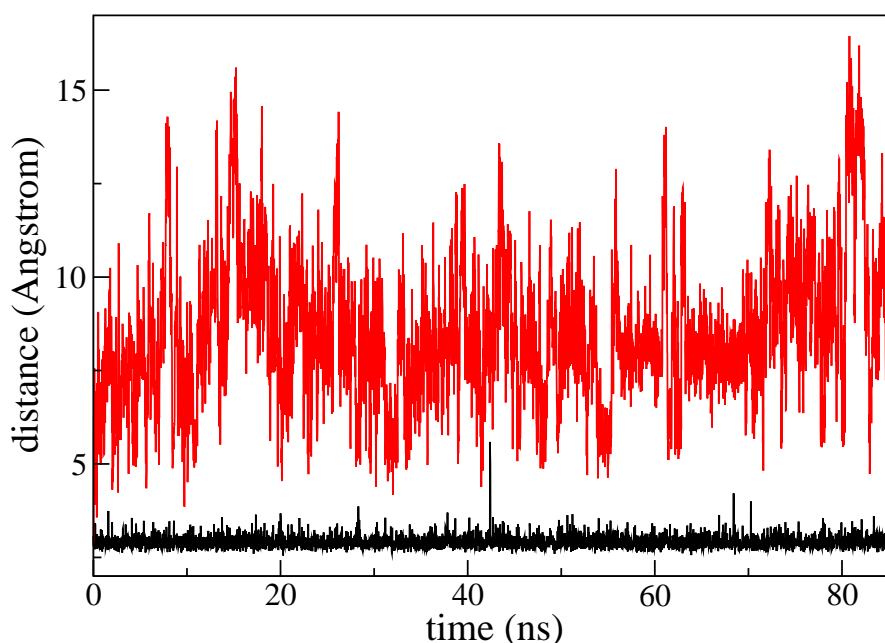


Figure 4.6: Comparison of Arg52 movement for MM MD with protonated (black) and neutral Arg52 (red) by measuring the distance between Arg52 NH and Tyr98 backbone O. For protonated Arg52 the distance between the Arg52 side chain and the Tyr98 backbone is constant at 2.75 Å in average. The hydrogen bond between both residues is intact throughout the simulation. This hydrogen bond does not exist for the simulation with deprotonated Arg52. Its side chain rotates away from Tyr98 into the bulk water and the hydrogen bond is broken.

Because of the broken hydrogen bond between Arg52 and Tyr98, a channel was opened through which bulk water molecules were able to enter the chromophore binding pocket. These water molecules hydrogen bonded to O1

for a short time, mostly less than 20 fs, and moved away into the bulk again. Up to two water molecules were found hydrogen bonded to O1 which is shown in Figure 4.7 for every 20th step of the simulation. The following picture, Figure 4.8, shows exemplarily chosen water molecules hydrogen bonded to O1. The graphs depict the distance between O1 and the respective water molecule during the MM MD simulation.

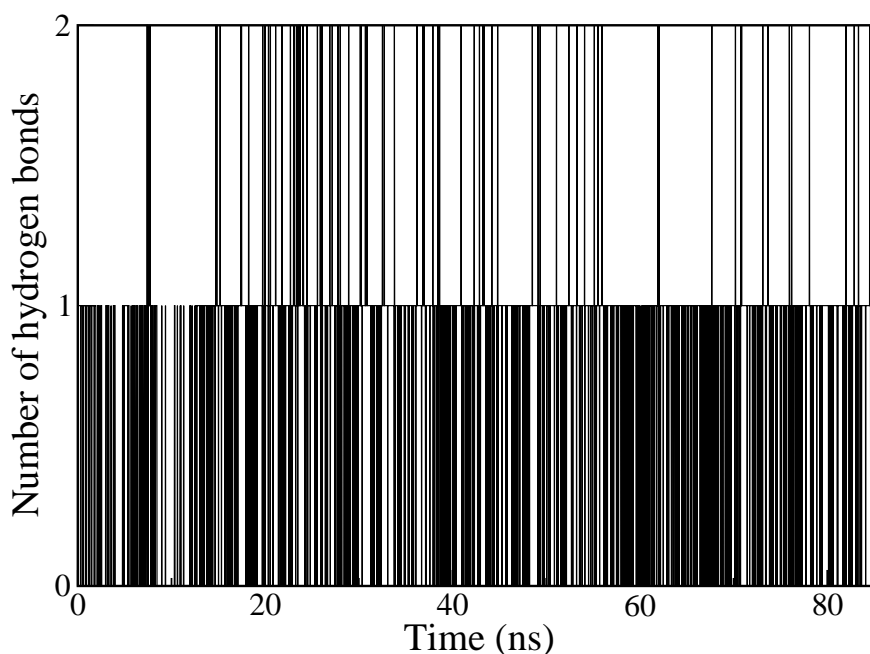


Figure 4.7: The number of hydrogen bonds between the chromophore's O1 and water molecules from the bulk for every 20th step of the simulation of PYP in water with neutral Arg52. Water molecules are able to enter the chromophore binding pocket because of the Arg52 side chain movement. Up to two water molecules are found bound to O1 at a time.

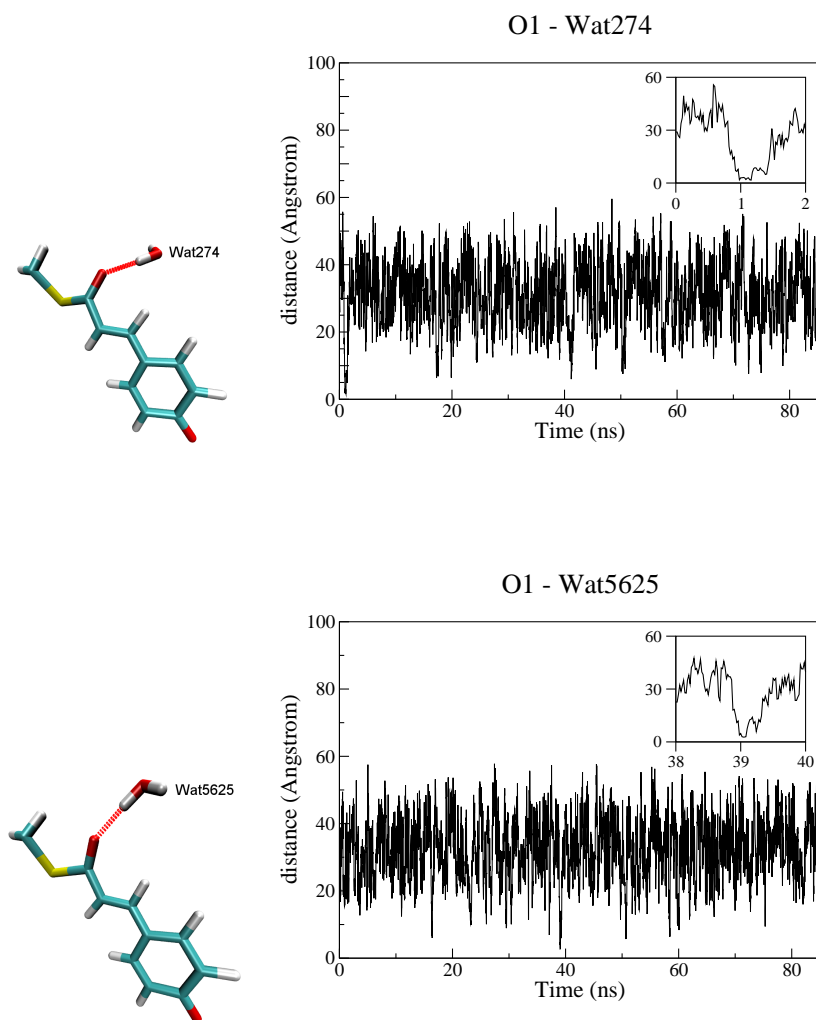


Figure 4.8: Water molecules from the bulk are able to enter the chromophore binding pocket and form hydrogen bonds to the chromophore O1. These hydrogen bonds are stable for mostly less than ten steps whereupon the water molecules move away into the bulk again. Four water molecules (Wat274, wat285, wat3781 and wat5625) were randomly chosen to illustrate this occurrences. This figure shows the water molecules wat274 and wat5625 hydrogen bonded to the chromophore O1 and the distances these water molecules have during the MM MD simulation to the O1. The data for the remaining two water molecules is shown in the appendix (figure 7.2).

One water molecule, which entered the protein following the pathway from the bulk through the channel opened by the movement of the Arg52 side chain into the chromophore binding pocket, took a position near the chromophore where it bridged Thr50 and the phenyl oxygen of the chromophore. Thr50 underwent a considerable movement as well upon movement of the Arg52 side chain because the hydrogen bond formed between these two residues in the initial structure was broken. As a result, it was not able to form a hydrogen bond to the chromophore without the bridging water molecule. Additionally, the Tyr98 side chain moved away from the position observed in the X-ray structure. As the Arg52 side chain and the Tyr98 back bone form a hydrogen bond in the X-ray structure, it is probable that this movement is due to the fact that the neutral Arg52 could not form a stable hydrogen bond to Tyr98 anymore and thus both amino acid side chains moved into a more favourable position. The positions resulting from these movements are depicted in Figure 4.9.

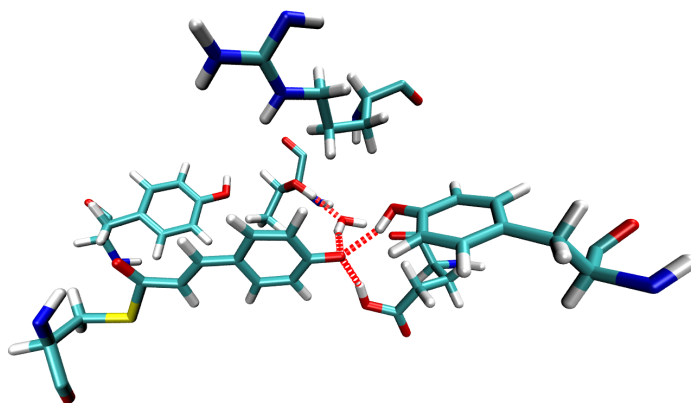


Figure 4.9: Structure of the chromophore binding pocket and surrounding amino acids after the hydrogen bond between Arg52 and Tyr98 has been broken and side chains of both amino acids moved into their new position. The water molecule bridging Thr50 and the chromophore was able to enter the binding pocket because of the before described movement.

In one simulation, the chromophore moved away from its initial position in the binding pocket in direction of the bulk water for a short time of 600 fs. During this event, the hydrogen bonds to Glu46 and Tyr42 were broken.

An additional water molecule entered the chromophore binding pocket and bridged Glu46 and Tyr42 to O4a of the chromophore, which is shown in Figure 4.10. The structure of the hydrogen bonded network of the chromophore was re-established after the chromophore moved back and resumed its initial position. This conformation remained stable for the successive part of the simulation. Figure 4.11 depicts the distance between the chromophore O4a and the respective oxygen atoms of Tyr42 and Glu46 which participate in hydrogen bonds to O4a in the initial structure.

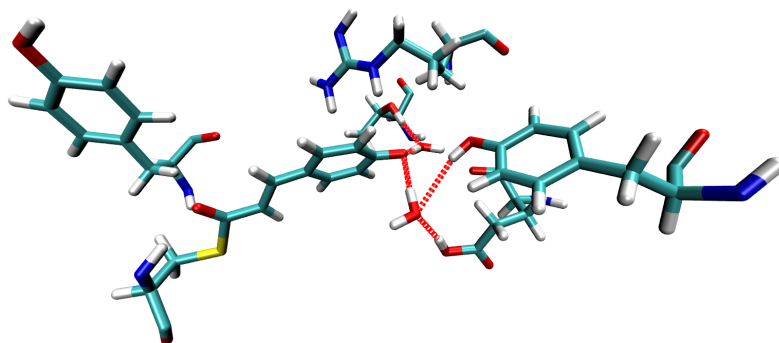


Figure 4.10: Chromophore binding pocket during the movement of the chromophore. The hydrogen bonds between the chromophore and Tyr42 and Glu46 are preserved by a bridging water molecule which was able to enter the chromophore binding pocket because of the broken hydrogen bond between Arg52 and Tyr98.

Comparing the RMS fluctuations (RMSF) of the MM MD simulations for both Arg52 protonation states in Figure 4.12 shows that the RMSF differ for Arg52 and Tyr98 and neighbouring amino acids. The fluctuations of the remaining amino acids of PYP are alike for both Arg52 protonation states. The large RMSF value of residue 114, a Serine positioned at the surface of PYP, observed for both Arg52 protonation states, is due to the fact that its side chain rotates freely in the bulk water and does not form permanent hydrogen bonds to other parts of PYP.

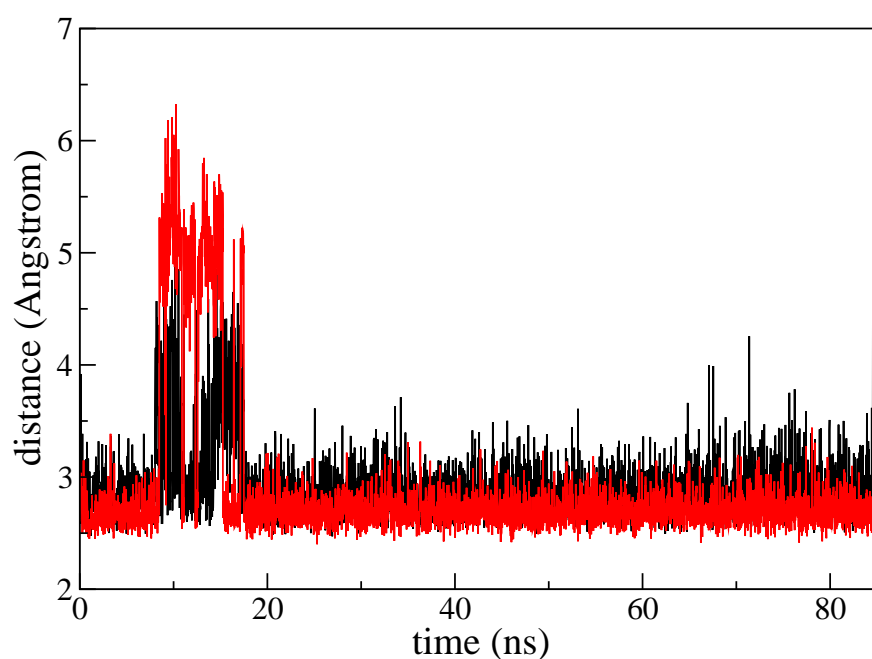


Figure 4.11: Length of hydrogen bonds between the chromophore and Tyr42 (black) and Glu46 (red) during MM MD simulations with deprotonated Arg52. The hydrogen bonds are broken between step 430 and 730 because the chromophore moves out of the binding pocket. Afterwards the initial conformation is re-established for the remaining simulation time.

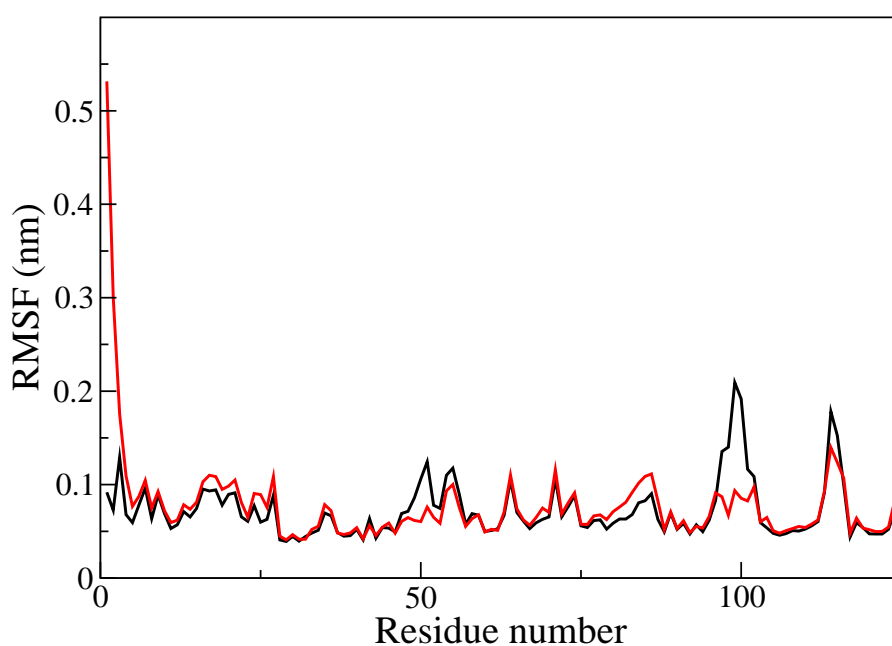


Figure 4.12: Comparison of the RMSF values of the individual PYP residues for protonated (red) and neutral Arg52 (black). Residues 52 and 98 and the neighbouring amino acids show the largest difference between both protonation states due to the large displacements observed in the MM MD simulations with neutral Arg52. Fluctuations of Ser114 are observed because this residue, located at the protein surface, does not form hydrogen bonds to other parts of the protein and therefore its side chain is able to move freely.

4.1.1 QMMM Dynamics

The MM MD simulations of PYP in water with neutral and protonated Arg52, which were presented in the preceding sections, showed hydrogen bond lengths between the chromophore and Tyr42 and Glu46 that differed from the X-ray structure used as initial structure. These hydrogen bonds were 2.52 and 2.56 Å long in the X-ray structure and had average lengths of 2.8 and 2.7 Å in the MM MD simulations.

The reproduction of the hydrogen bond length published in the X-ray structure 2ZOI would help to discover the effect of short hydrogen bonds on the dynamics. Therefore, QMMM MD simulations were used to observe whether the hydrogen bond lengths can be conserved in dynamics.

For the ground state QMMM MD simulations, the same initial setup as for the ground state MM MD simulations was used with the chromophore, Cys69, Tyr42 and Glu46 forming the QM region. 5000 steps of 0.001 ps starting from the minimised structures were performed with B3LYP/6-31G* as QM method and the Amber03 force field and the TIP3P water model for the MM part. The protein structure was fitted to the initial structure at each step.

The hydrogen bond lengths shown in Figure 4.13 have an average of 2.7 Å for the hydrogen bond to Tyr42 and 2.6 Å for the hydrogen bond to Glu46, which is longer than in the X-ray structure where these hydrogen bonds have a length of 2.52 and 2.56 Å respectively. Compared to MM MD simulations where 2.8 and 2.7 Å were observed, respectively, the hydrogen bonds are closer to the X-ray structure values.

Figure 4.14 shows the movement of the side chain of the neutral Arg52 by the length of the hydrogen bond from NH2 of Arg52 to the backbone O of Tyr98. The side chain did not move into the bulk as observed in the MM MD simulation but shows larger movement compared to the QMMM MD simulation with protonated Arg52. These movements include breaking of the hydrogen bond between Arg52 and Tyr98 but also reformation of this bond.

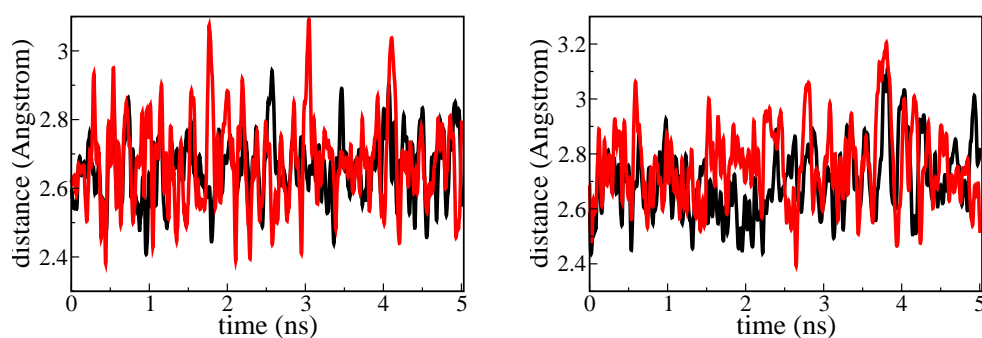


Figure 4.13: Hydrogen bond length between (A) the chromophore's O4a and Glu46 and (B) the chromophore's O4a and Tyr42 for QMMM MD simulations with protonated (black) and neutral Arg52 (red). For both protonated and neutral Arg52 the hydrogen bond length is in average the same.

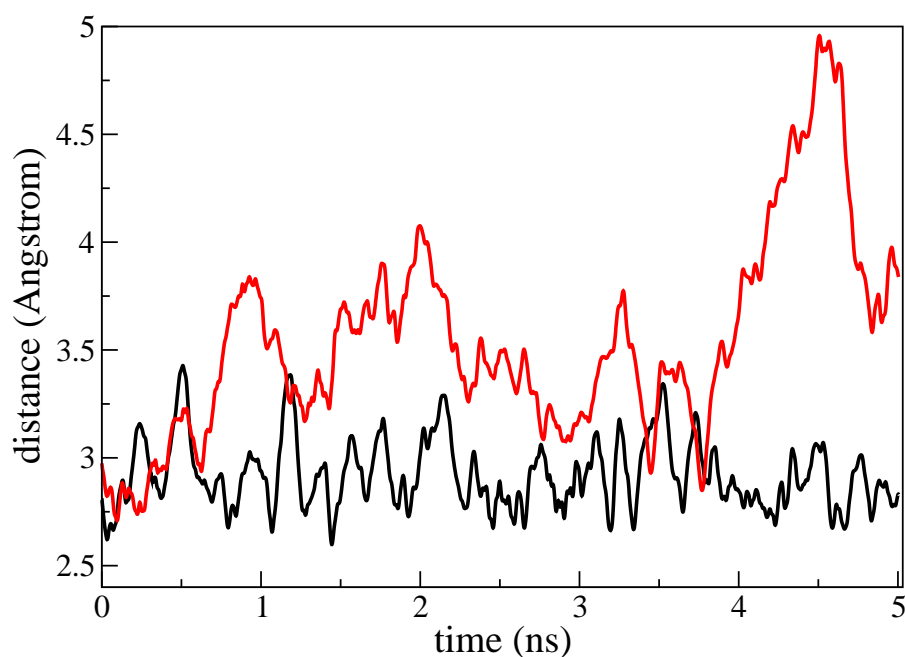


Figure 4.14: Distance between NH of Arg52 and the backbone O of Tyr98 for QMMM MD simulations with protonated (+) and neutral (x) Arg52. For neutral Arg52 the distance between the Arg52 side chain and the Tyr98 backbone is larger than for protonated Arg52. The hydrogen bond existing in case of protonated Arg52 is not formed in case of neutral Arg52.

The effect of using diffuse functions during QMMM dynamics on the hydrogen bond length between O4a and Tyr42 and Glu46 has been investigated by two simulations starting from the same initial structure for each, protonated and neutral Arg52, and were carried out for 500 steps of 0.001 ps. The MM part of the system was described by the Amber03 force field with the TIP3P water model and the QM methods were B3LYP/6-31G* and B3LYP/6-31+G*.

Both for protonated and neutral Arg52, the length of the hydrogen bonds did not differ significantly during the simulation time. The evolution of the bond length over time is depicted in Figures 4.15 and 4.16. The usage of diffuse functions did not show a trend towards either shorter or longer hydrogen bonds.

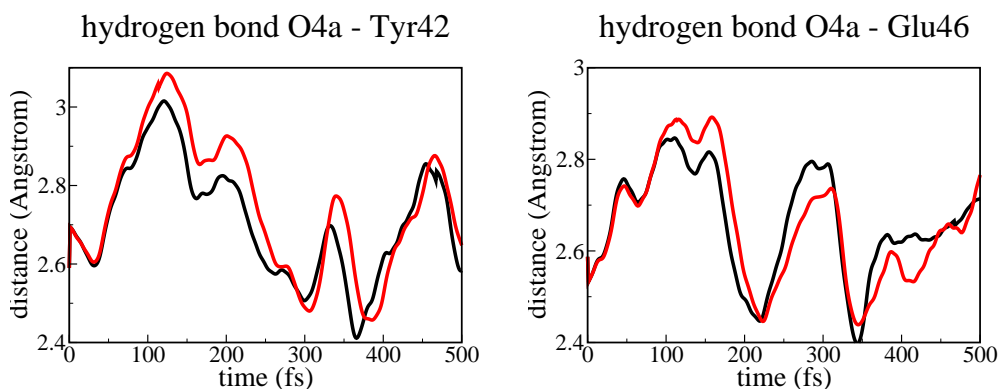


Figure 4.15: Evolution of the hydrogen bond length between O4a and Tyr42 (left) and Glu46 (right) during QMMM MD simulations with (red) and without (black) diffuse functions for PYP with protonated Arg52. The usage of diffuse functions did not provoke shorter or longer hydrogen bonds compared to not using diffuse functions.

As in the ground state MM MD simulations, formation of the hydrogen bond between O4a and Thr50 could be observed in ground state QMMM MD simulations with both protonated and neutral Arg52. The different description of the QM region in both ground state MD simulation setups did not introduce different behaviour in this case.

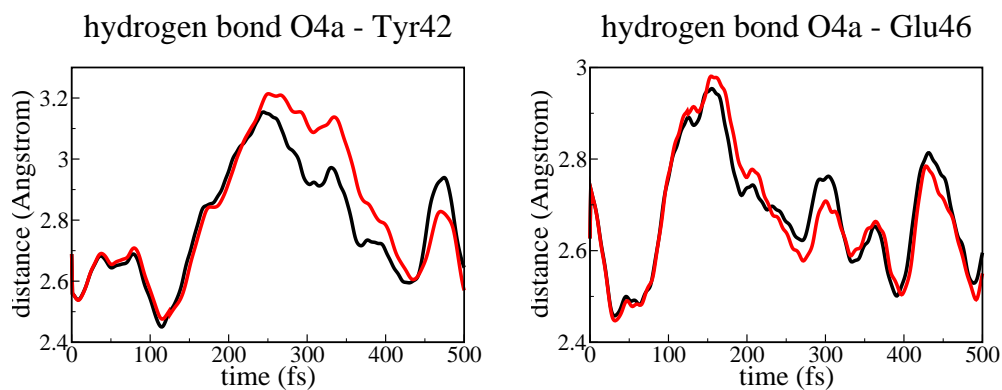


Figure 4.16: Evolution of the hydrogen bond length between O4a and Tyr42 (left) and Glu46 (right) during QMMM MD simulations with (red) and without (black) diffuse functions for PYP with neutral Arg52. The usage of diffuse functions did not provoke shorter or longer hydrogen bonds compared to not using diffuse functions.

4.2 Discussion Ground State MD Simulations

The performed MD simulations of PYP in water showed that the protonation state of Arg52 influences the protein dynamics. The deprotonation of Arg52 leads to large configurational changes in direct vicinity of the chromophore. The most prominent observation upon deprotonation of Arg52 is the movement of the Arg52 and the Tyr98 side chain. They are no longer in the position that is assigned to it by the crystal structures but moved away from each other. Both form hydrogen bonds to multiple amino acids and show large mobility compared to simulations with protonated Arg52.

This motion is similar to observations described for the pB state of PYP. Arg52 is ascribed the role of a lid on the chromophore binding pocket, which has been described in several publications.

In their review Hellingwerf *et al.*¹⁰⁴ depict rearrangements of the chromophore binding pocket upon formation of the signalling state pB. These rearrangements include breaking of the hydrogen bond between Arg52 and Tyr98 and exposure of the chromophore to the solvent.

Shimizu *et al.*²¹ propose that Arg52 acts as a lid on the chromophore binding pocket by forming hydrogen bonds to Tyr98 and Thr50. This function prevents water molecules from entering the binding pocket. In their study on the structure of the R52Q mutant they found it structurally mimics the pB structure in the wild type. Therefore, they propose that the Arg52 side chain moves away in the pB state and thus opens the chromophore binding pocket. The accessibility of the chromophore to water molecules helps to change its pK_a to a value where it can be protonated by Glu46.

In accordance with the above described observations, Imamoto *et al.*¹⁰ reflect in their review on experimental studies, which observed changes in the binding of organic compounds to PYP upon formation of the signalling state pB. Organic anions bind to PYP in its signalling state but not in the ground state, indicating that a positively charged residue is exposed to the solvent in the signalling state²⁰. The same is found for the reactivity to lipids and hydrophobic agents which react with hydrophobic parts of PYP exposed in the signalling state^{105,106}. They concluded that these observations agree with the movement of the Arg52 side chain and subsequent opening of the chromophore binding pocket.

The behaviour of the Arg52 side chain presented for neutral Arg52 in this thesis is very similar to the behaviour of this residue in the pB state of PYP described by the aforementioned reviews and studies with the difference that Arg52 is assumed to be protonated in the wild type pR state. Neutral Arg52 acts like a lid on the chromophore binding pocket and by rotating into the solvent allows water molecules to enter the cavity. This movement happens

because the hydrogen bonds formed in the wild type by the protonated Arg52 with the amino acids Thr50, Val66 and Tyr98 to close the pathway into the chromophore binding pocket are broken upon deprotonation of Arg52. In the wild type pR state, the Arg52 movement is caused by the isomerised and protonated chromophore. The conformational change is functional in wild type as it refers the signal that absorption of blue light occurred to the next protein in the signalling cascade.

A major difference between the structures from MM simulations presented in this thesis and the crystal structure of Yamaguchi *et al.*, for both protonated and neutral Arg52, is the length of the hydrogen bonds formed by O4a to Tyr42 and Glu46. In the simulations, these hydrogen bonds are in average 2.8 and 2.7 Å long, respectively, because the force field used for the MM part is parametrised to reproduce a hydrogen bond distance of about 2.7 Å. QMMM MD simulations show shorter hydrogen bonds, 2.7 and 2.6 Å, respectively, than MM MD simulations. The extremely short hydrogen bond length of the crystal structure can not be reproduced by either method though using QMMM is better suited than MM.

The QM method used during the QMMM MD simulations was B3LYP/6-31G*. Previous publications on the accuracy of the description of hydrogen bonds showed that B3LYP is able to predict the geometry of hydrogen bonds better than Hartree Fock^{107,108}. Both methods with different basis sets were compared to second-order Møller-Plesset perturbation theory (MP2), which is a high accuracy ab initio method that is computationally expensive and its use therefore mostly restricted to small systems. This comparison of B3LYP to MP2 revealed that usage of diffuse functions enhances the description of hydrogen bonds¹⁰⁷. The usage of B3LYP with an appropriate basis set is suggested in these publications when MP2 can not be afforded due to computational limitations. Because of this B3LYP has been used for testing semi-empiric methods, for example SCC-DFTB^{109–111}, on systems too large for MP2 calculations due to computational resources.

The QMMM MD simulations of PYP in water in this thesis were carried out without usage of diffuse functions. For both protonation states the effect of diffuse functions on the length of the hydrogen bonds between O4a and Tyr42 and Glu46 was examined. The differences were less than 0.1 Å compared to simulations without diffuse functions and no trend either to longer or shorter hydrogen bond length could be observed. This observation justifies the use of the computationally less expensive method without diffuse functions.

In the preceding chapter, we found that the delocalised proton of the hydrogen bond between the chromophore and Glu46 stabilises the crystallised

protein with neutral Arg52 by 3.6 kJ/mol. It is possible that a comparable stabilisation occurs in PYP in water. But this energy gain would not be enough to compensate the energy losses from the broken hydrogen bonds between Arg52 and Thr50 and Tyr98. The delocalisation would not stabilise the position of the Arg52 side chain because no additional electrostatic attraction would arise from it.

Simulations with both protonated and neutral Arg52 show hydrogen bonds between Thr50 and O4a, either direct or bridged by a water molecule. This observation is in contrast to X-ray structures where Thr50 forms a hydrogen bond to Tyr42^{18,22,49,89}. Gromov *et al.*³⁶ conjecture in their theoretical study on the existence of a direct hydrogen bond between Thr50 and O4a. They calculated the energy gaps between the ground state and several excited states for different model systems of the chromophore that included a varying composition of the amino acids surrounding the chromophore. The energy gap between the ground state and the first excited state was lowered by 0.05 eV after modifying the Thr50 side chain such that it formed a hydrogen bond to O4a. Further QMMM calculations supported this observation. A hydrogen bond between Thr50 and O4a was experimentally observed in the Y42F mutant^{19,93}. In this structure, Thr50 provided the second hydrogen bond that could not be formed otherwise due to the mutation.

The formation of the hydrogen bond between O4a and Thr50 in this thesis is a possible consequence of the elongation of the short hydrogen bond between O4a and Glu46. The additional stabilisation by the third hydrogen bond may be needed to stabilise the negative charge on O4a when the SSHB is absent, which is the case in our simulations because the hydrogen bond was described by the Amber03 force field which is parametrised for a hydrogen bond length of 2.7 Å. Formation of the hydrogen bond was observed for both MM and QMMM ground state MD simulations with protonated and neutral Arg52. Thr50 was described by the Amber03 force field in all simulations. This fact stimulates to consider the used force field as source of formation of this hydrogen bond. Re-evaluating the simulation protocol by finding a QM method which reproduces the SSHB between O4a and Glu46, using different force fields or including Tyr42, Glu46 and Thr50 into the QM region would help to shed light on this matter.

In QMMM MD simulations in this thesis, the movement of the neutral Arg52 side chain is not as pronounced as in the MM MD simulations. The hydrogen bond to Tyr98 is broken but the side chain does not move away as far from its initial position as in the MM MD simulations. The large movement of the Arg52 side chain could be an artefact of the MM description

of the protein. Or the description in case of the QMMM simulations could be unbalanced which may lead to a too strong stabilisation of the Arg52 side chain in its initial position. Another possible source for the observed difference is the shorter timescale of the QMMM MD simulations. Further simulations are needed to look into this topic.

Nakamura *et al.*¹¹² studied the PYP wild type and the E46Q mutant by Raman spectroscopy. Their interest laid in the development of the hydrogen bonded network of the chromophore throughout the photocycle. By comparing the spectra of the wild type and the mutant they observed differences in the ground state for a band at 1555 cm^{-1} which is sensitive to the hydrogen bonds of O4a. The relative intensity for the wild type was higher than for the mutant. Upon excitation the intensity of this band for the wild type decreased immediately to the intensity observed for the mutant within 150 fs. This observation was interpreted as rearrangement of the hydrogen bonded network from a short hydrogen bond between O4a and Glu46 in the ground state to a longer hydrogen bond, as it exists in the mutant, in the excited state.

This conclusion would mean that the charge migration, which is initiated by the excitation, is followed immediately by the elongation of the hydrogen bond between O4a and Glu46. In the light of this observation, the elongation of the hydrogen bonds during MD simulations is still unfortunate because the used methods can not reproduce them but the results of the excited state QMMM MD simulations presented in the following section are not biased by the hydrogen bond lengths.

Deprotonation of Arg52 destabilises the protein structure in direct vicinity of the chromophore binding pocket in the simulations presented in this thesis. The broken hydrogen bond between Arg52 and Tyr98 induces movements of these two amino acids, opening a channel for bulk water molecules towards the chromophore. On the level of theory used, these observations lead to the conclusion that the structure of Yamaguchi *et al.* with neutral Arg52 is not stable for the soluted protein. The stability of the structure in a crystal can not be evaluated by this work.

4.3 Deactivation Events

The structural progression of PYP in the ground state observed in this thesis differs for both considered Arg52 protonation states. The influence of these differences on the deactivation events from the first excited state to the ground state will be described in the following section.

For the excited state QMMM MD simulations 85 (protonated Arg52)/87 (neutral Arg52) snapshots were taken from the respective ground state MM MD simulations described above. Starting from these snapshots, ground state QMMM MD simulations with CASSCF 8,7/6-31G* as QM method were set up for 1000 steps of 0.001 ps. Following these ground state simulations, excited state simulations were started using the same QM method and time step. All QMMM calculations were carried out using the GROMACS/GAUSSIAN03 interface. The MM part was described with the Amber03 force field using the TIP3P water model. To allow for hops from the excited state to the ground state at the conical intersection, the surface hopping algorithm developed by Groenhof *et al.* described in section 2.6.1 was applied.

The majority of QMMM MD simulations showed no deactivation from the excited state during the simulation time. From these simulations the ones with protonated Arg52 have mainly db twisted conformation of the chromophore, 50 simulations, while simulations with neutral Arg52 have sb twisted conformation, 30 simulations. The respective other twisted conformation as well as a planar conformation is observed only in few simulations as can be seen in Table 4.1.

For all simulations which show surface hops during the simulation time of 3 ps, double bond (db) isomerisation of the chromophore happened more often for the system with protonated Arg52 (11.4%) than for the system with neutral Arg52 (9.3%). PYP with neutral Arg52 deactivated mainly via single bond (sb) rotation (43 %).

The average times for deactivations from the excited state to take place are shown in table 4.2. They are shorter for db isomerisation to happen than for sb rotation. The time until deactivation from excited state took place was shorter overall for the system with neutral Arg52.

Four trajectories for protonated Arg52 and one trajectory for neutral Arg52 showed deactivation from the excited state via isomerisation around the double bond and subsequent rotation on the ground state to the *cis*-

Table 4.1: Observed torsions for excited states QMMM MD simulations. The values for simulations which showed deactivation during the simulation time are written in brackets. Simulations with neutral Arg52 show more deactivations than simulations with protonated Arg52. In the latter case proportionally more deactivations took place via db isomerisation. Neutral Arg52 favoured deactivation via sb rotation.

Arg52 protonation state	observed torsion			ratio (of all runs) in %		
	db twist	sb twist	planar	db	sb	planar
protonated	60 (10)	20 (16)	7	69.0 (11.4)	23.0 (18.2)	8.0
neutral	15 (8)	67 (37)	3	17.6 (9.3)	78.8 (43.0)	3.6

Table 4.2: Average time until deactivation via both pathways from the excited state for different Arg52 protonation states. The deactivation via db isomerisation is faster than via sb rotation. Deactivations happen earlier for the case of neutral Arg52.

Protonation state	steps til db isomerisation	steps til sb rotation
protonated	1346	1892
neutral	1222	1470

conformation. The hydrogen bond between O1 of the chromophore and the Cys69 backbone remains intact throughout the entire process. The intact hydrogen bond induces a strained geometry of the chromophore as the ethylene chain can not adopt a favourable flat conformation but is twisted to maintain the hydrogen bond. This twist of nearly 90 degree between the C1-O1 bond and the phenyl ring is distributed over the three dihedral angles in-between. The double bond is nearly in *cis*-conformation as it is bent by 30 degrees out of the plane to make the twist possible while the other dihedrals are twisted by about the same value. This conformation is adopted by the chromophore in 3 trajectories and is shown in Figure 4.17-1. In the remaining 2 trajectories the double bond shows no bent and the 90 degree twist is compensated solely by the dihedrals over the two adjacent bonds as shown in Figure 4.17-2.

All five QMMM MD simulations continued on the ground state after deactivation from the first excited state. In 70 to 525 steps all simulations maintained the twisted *cis*-conformation.

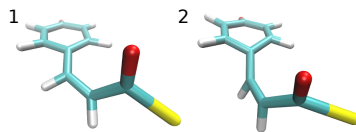


Figure 4.17: Depiction of strained geometry of the chromophore's ethylene chain after deactivation from excited state to ground state in *cis*-conformation. (1) shows the even distribution over the three affected bonds whereas the twist is not dispersed over the double bond in (2).

These five simulations, which deactivated via db isomerisation to the ground state *cis*-conformation, were simulated further on ground state. B3LYP/6-31G* was used as QM method while the remaining part of the system was simulated using the AMBER03 force field and TIP3P water model. These ground state simulations were carried out for 10000 steps of 0.001 ps each. During the simulations one immediate re-isomerisation to the *trans*-conformation was observed while the remaining four simulations retained the *cis*-conformation. Of these four simulations, a breakage of the hydrogen bond between O1 and the Cys69 backbone was observed for two simulations within 0.75 and 1.0 ps. The hydrogen bond stayed intact for the simulation time in the two remaining simulations (Table 4.3). The structures of the chromophore for intact and broken hydrogen bond between O1 and Cys69 N are depicted in Figure 4.18

Table 4.3: Events on ground state after deactivations from the excited state via db isomerisation. Of five simulations which deactivated to the *cis*-conformation on the ground state four maintain this conformation. For two of these the hydrogen bond to Cys69 is broken during the ground state QMMM MD simulations.

protonation state	run number	conformation	hydrogen bond to Cys69
neutral	13	<i>cis</i>	broken
protonated	1	<i>cis</i>	intact
	28	<i>cis</i>	broken
	48	<i>trans</i>	intact
	67	<i>cis</i>	intact

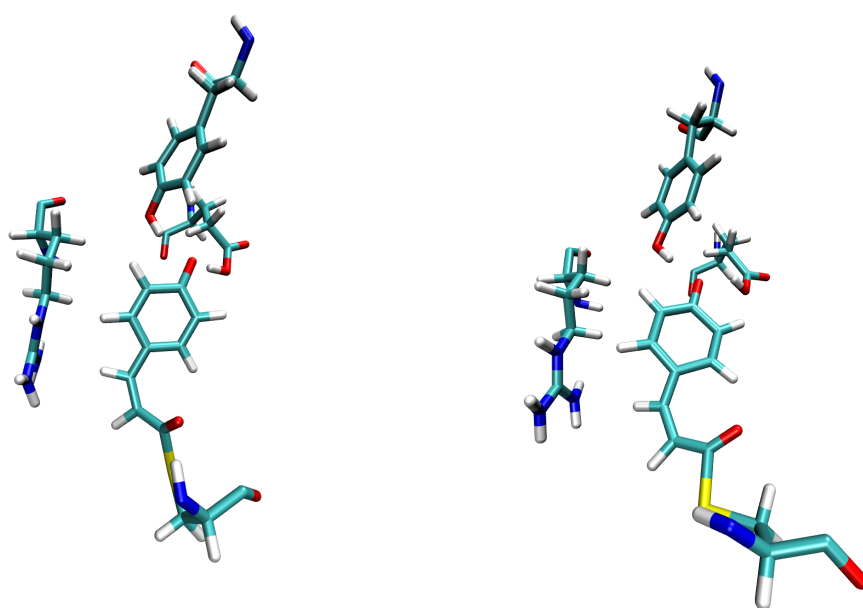


Figure 4.18: Depiction of the hydrogen bond between the chromophore O1 and the Cys69 N after deactivation from the excited state to *cis*-conformation on the ground state and subsequent QMMM MD simulation. The left picture shows the intact hydrogen bond and a strained geometry of the chromophore while the picture on the right displays the broken hydrogen bond with the unstrained chromophore.

The observed deactivation events and the actual situation of the hydrogen bonded network at the time of excitation, i.e. the beginning of the excited state QMMM MD simulations, are correlated. For the latter the number of existing hydrogen bonds and their length were analysed by using the Gromacs tool `g_hbond`.

For both protonated and neutral Arg52, structures from which trajectories with db isomerisation started show on average more hydrogen bonds to O4a and less hydrogen bonds to O1 than structures from which trajectories with sb rotation resulted. These differences are larger for the hydrogen bonds to O4a in the case of protonated Arg52 and for hydrogen bonds to O1 in case of neutral Arg52. The average hydrogen bond counts are outlined in table 4.4. Tables 4.5 and 4.6 comprise the hydrogen bond count to the oxygen atoms for both Arg52 protonation states.

Including the simulations with twisted chromophore that have not decayed within the simulation time into the analysis of the correlation between average hydrogen bond number at the moment of excitation and deactivation event, the trend described above is conserved.

Table 4.4: Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for neutral and protonated Arg52

Arg52 protonation state	deactivation event	average number of HBs to	
		O1	O4a
protonated	db	1.00	2.90
	sb	1.125	2.3125
neutral	db	1.125	2.625
	sb	1.657	2.571

As in the ground state simulations with protonated Arg52, Thr50 is in hydrogen bond distance to the chromophore's O4a occasionally. Hydrogen bonds between Thr50 and Tyr42 could not be observed during the simulations. Thus, in this theoretical study Thr50 does stabilise the negative charge on the chromophore directly and not indirectly via Tyr42 in the excited state as well as in the ground state. This observation is in disagreement with structural studies and may be a source for artefacts in the theoretical calculations presented.

The average hydrogen bond distance between the chromophore O1 and the Cys69 N is shorter at the moment of excitation for the trajectories resulting in db isomerisation than for the structures which yielded sb rotation

in the case of protonated Arg52 (sb 2.93, db 2.87). For neutral Arg52 the difference in the average distances is smaller (sb 2.94, db 2.91). This observation is due to the additional water molecules from the bulk which form a hydrogen bond to O1, which needs additional space and results in increased hydrogen bond lengths.

Water molecules were found hydrogen bonded to both O4a and O1 of the chromophore for the simulations with neutral Arg52 as they were able to enter the chromophore binding pocket in the preceding ground state MM MD simulations. One water molecule is stable near O4a and forms a bridge from Thr50 to the chromophore. This bridge provides the third possible hydrogen bond to O4a that can be observed in the setup used in this thesis. It would else be missing in the case of neutral Arg52 because of the structural changes caused by the altered protonation state. Two other water molecules form bridges from Tyr42 and Glu46 to the chromophore in two structures as the chromophore moved out of the binding pocket as described in section 4.1. The water molecules, which form hydrogen bonds to O1 additionally to the Cys69 N in several trajectories for the simulations with neutral Arg52, are the reason for the overall higher average number of hydrogen bonds to this atom.

Table 4.5: Number of hydrogen bonds at the moment of excitation for all individual simulations with neutral Arg52 that showed deactivation from the excited state to the ground state.

sb rotation			db isomerisation		
run number	count of hbs to		run number	count of hbs to	
	O1	O4a		O1	O4a
6	1	2	0	1	3
8	2	3	7	1	3
14	0	3	13	1	3
15	1	3	17	1	3
18	2	2	27	1	2
20	2	2	30	2	3
21	1	2	52	1	2
22	2	1	75	1	2
28	2	3			
29	2	3			
32	2	3			
33	2	2			
35	1	3			
36	2	2			
38	2	3			
39	0	3			
40	1	2			
42	2	3			
45	2	3			
47	2	3			
48	2	3			
53	2	3			
54	2	2			
56	1	3			
61	2	2			
63	2	3			
66	1	3			
69	2	2			
70	2	2			
71	2	3			
72	0	2			
73	2	2			
78	2	3			
79	2	3			
83	2	2			
85	2	3			

Table 4.6: Number of hydrogen bonds at the moment of excitation for all individual simulations with protonated Arg52 that showed deactivation from the excited state to the ground state.

sb rotation			db isomerisation		
run number	count of hbs to		run number	count of hbs to	
	O1	O4a		O1	O4a
0	1	2	1	1	3
2	1	2	6	1	3
4	2	2	12	1	3
5	1	2	25	1	3
8	1	3	28	1	3
19	1	3	48	1	3
20	1	2	52	1	2
35	1	2	66	1	3
36	1	3	67	1	3
44	1	3	75	1	3
45	1	3			
51	2	1			
61	1	2			
69	1	2			
76	1	3			
86	1	2			

Influence on Barrier Heights

In X-ray structures, three hydrogen bonds to surrounding amino acids are formed by the chromophore in the binding pocket, two between O4a and Tyr42 and Glu46, respectively, and one between O1 and the backbone of Cys69. These hydrogen bonds influence the possible isomerisation pathways of the chromophore as has been shown by previous theoretical studies^{26,27,29,33,36,113}. In these studies the following observation was presented: The hydrogen bonds to O4a support db isomerisation while the hydrogen bond to the backbone supports sb rotation.

This effect is due to the diverging distribution of the negative charge on the chromophore at different points of the photoisomerisation process. In ground state the negative charge is localised on the phenyl ring and accordingly stabilised by the two hydrogen bonds to this oxygen. Upon excitation the negative charge is shifted towards the thiol group of the chromophore and the hydrogen bond to O1 stabilises this process (Figure 4.19-1, -2). Therefrom, double and single bond characteristics become less pronounced, which enables the isomerisation around the double bond.

The formation of the double bond twisted conformation is supported by hydrogen bonds to O4a. These hydrogen bonds stabilise the negative charge, which is shifted back towards the phenyl ring in the db twisted conformation. The twist around the single bond is supported by the charge delocalisation as well. During the rotation of the phenyl group around the adjacent single bond the negative charge is translocated further towards the thiol group, which results in a charge-transfer-state. Therefore, a strong hydrogen bond from O1 to the Cys69 back bone will support this deactivation pathway.

These shifts of the charge distribution explain why different hydrogen bond constellations support different processes in the protein in theoretical studies^{26,27,33}.

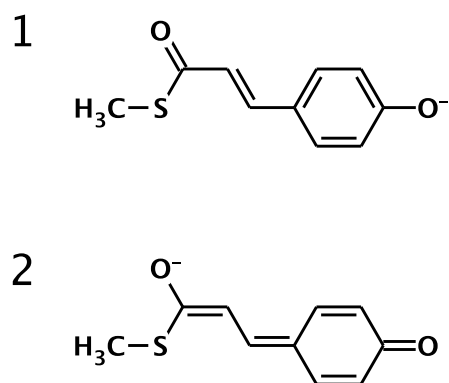


Figure 4.19: Charge localisation on the chromophore for ground state (1) and excited state (2). In ground state the negative charge is localised on O4a. Through light excitation a charge transfer towards the ethylene chain and O1 occurs.

The barrier heights for going from the planar intermediate to the db and sb twisted intermediates were calculated at different lengths for the hydrogen bonds between O4a and Tyr42 and Glu46. These hydrogen bonds are of particular interest due to their short length in the X-ray structure published by Yamaguchi *et al.*. The model system depicted in Figure 4.20 consisted of the chromophore, Tyr42, Glu46 and Cys69. By using the rigid scan function implemented in Gaussian09, the energy landscapes for 90 degree rotations around the single and double bond were recorded. The results of the rigid scans are the upper bounds for the energy barriers from the planar to the twisted structures. Using minimisation at each step as in a relaxed scan would lower the barriers and give better picture. Unfortunately problems with the convergence of these scans could not be resolved. The rigid scans were carried out for three different sets of hydrogen bond length to Tyr42 and Glu46:

- lengths taken from the X-ray structure (PDB entry 2ZOH), 2.52 and 2.56 Å respectively
- both hydrogen bonds at 2.70 Å
- both hydrogen bonds at 2.90 Å

The calculations were carried out using CAS 12,11/6-31G* as method and state averaging with equal distribution of the ground and first excited state. Convergence on wavefunction was achieved when deviations of the energy from the previous step were less than 10^{-6} Hartree within 512 steps.

The more the hydrogen bond lengths decrease the lower the db isomerisation and the higher the sb rotation barrier becomes (table 4.7).

Table 4.7: Comparison of the estimated barrier heights for single and double bond isomerisation for different hydrogen bond length. The shorter the hydrogen bonds to Tyr42 and Glu46 are, the higher the sb barrier and the lower the db barrier.

distance	db-barrier (kJ/mol)	sb-barrier (kJ/mol)
X-ray (2.52 and 2.56 Å)	21.40	22.45
2.70Å	24.26	17.12
2.90Å	27.41	13.00

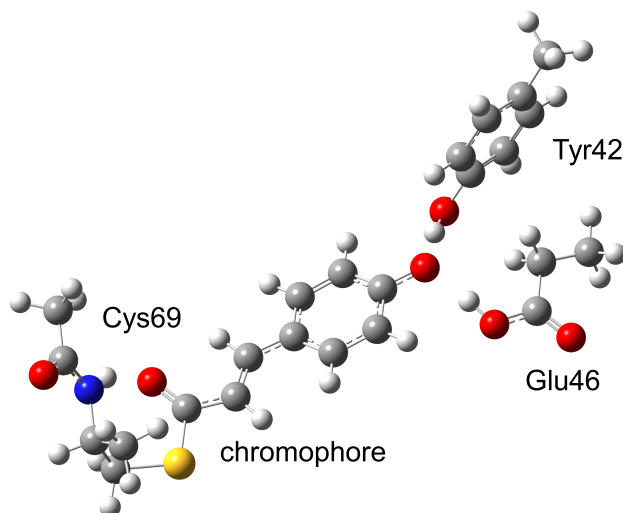


Figure 4.20: Chromophore with the amino acids Tyr42, Glu46 and Cys69

The barriers for both hydrogen bonds at 2.90 Å are shown in Figure 4.21. To reach the db twisted minimum from the planar conformation a barrier of 27.41 kJ/mol has to be crossed while the barrier to the sb twisted minimum is considerably lower at 13.00 kJ/mol.

When both hydrogen bonds are shortened to their respective length from the X-ray structure the barrier for sb rotation rises to 22.45 kJ/mol. In return, the barrier for db isomerisation is lowered to 21.40 kJ/mol as can be seen in Figure 4.22.

Steric interactions between the thiol group and the phenyl group of the chromophore arising from the use of the rigid scan method are equal for all calculations of the db twist as well as for all calculations of the sb twist. The trend for shortening of the hydrogen bonds for each kind of twist is not influenced by these interactions. The steric interactions differ between db and sb twist. Therefore, a direct comparison between the barriers of both kinds of twist must not be meaningful.

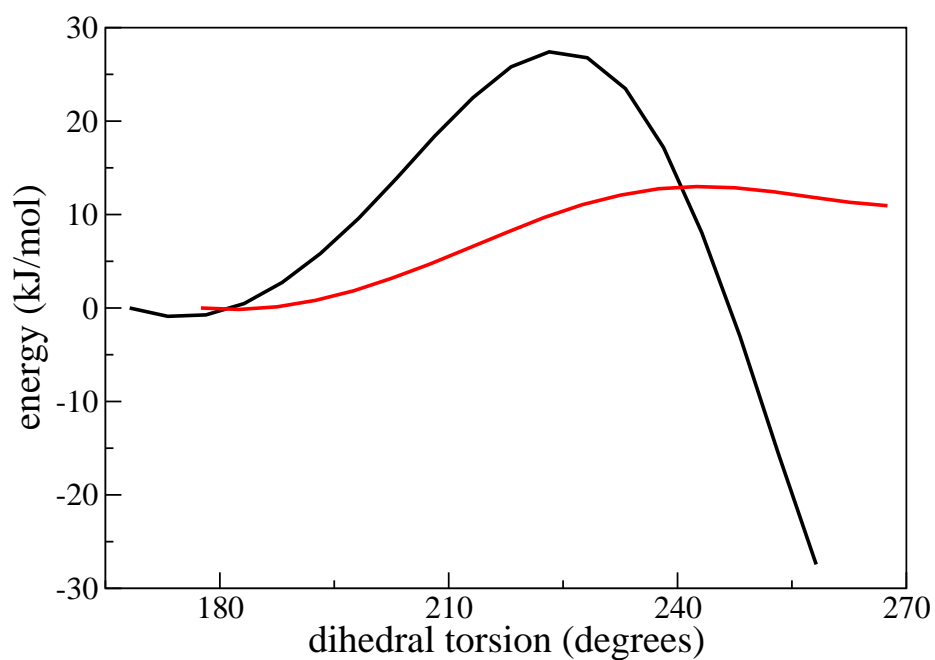


Figure 4.21: Single (red) and double bond (black) isomerisation barrier heights for hydrogen bond length of 2.90\AA each. The db barrier is significantly higher than the sb barrier.

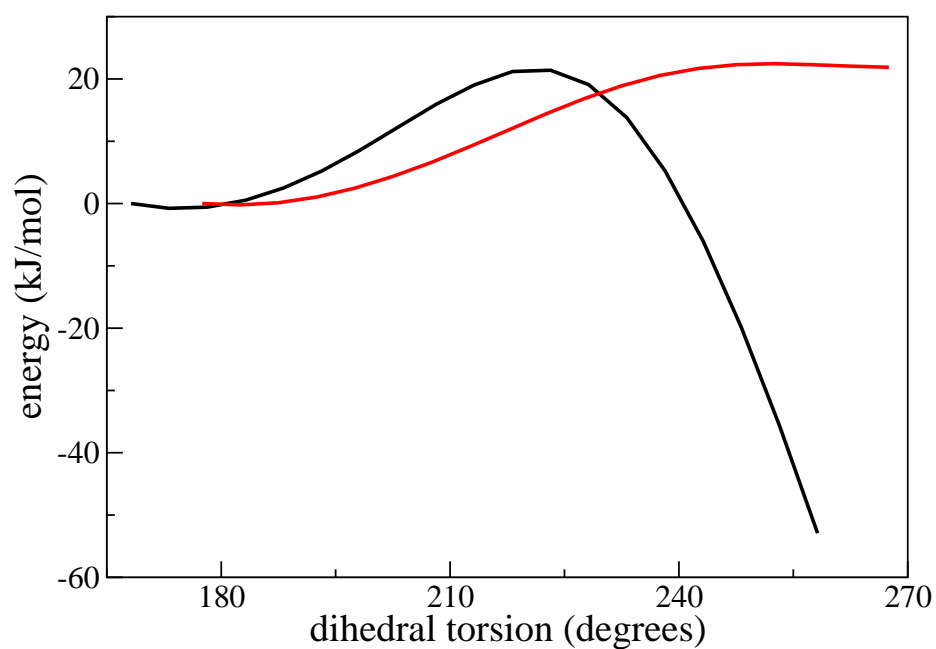


Figure 4.22: Single (re) and double bond (black) isomerisation barrier heights for hydrogen bond length taken from X-ray structure 2ZOH. Both barriers are of about the same heights.

4.4 Discussion of Excited State Deactivations

For the present investigation, deactivations from the excited state for both Arg52 protonation states take place via db isomerisation and sb rotation. A previous study of Groenhof *et al.*²⁵ on PYP in water with 14 trajectories and CAS 6,6/3-21G as QM method and the Gromos96 force field observed only deactivation via db isomerisation, which continued to both *cis*- and *trans*-conformation on the ground state. The difference between the mentioned study and the simulations presented in this thesis may be due to the different QM and MM methods. The influence of the QM method used on the excited state potential energy surface has been investigated with a model system introduced in chapter 2.5.

In the following, the discussion will be separated into two parts. The first will focus on the differences between sb rotation and db isomerisation while the second part will concentrate on the simulations that showed deactivation via db isomerisation.

At the moment of excitation, the average number of hydrogen bonds to O4a and O1 differs corresponding to the observed deactivation mechanism. Simulations that show deactivation via db isomerisation show more hydrogen bonds to O4a at the time of excitation than simulations that deactivate via sb rotation. For hydrogen bonds to O1 the opposite correlation is observed. This correlation is in agreement with a theoretical study²⁶, which determined that the negative charge is translocated upon excitation. In the ground state the charge is located on the phenyl ring while the negative charge in the excited state is predominantly found on the ethylene chain. Based on this study further theoretical studies^{27,29,33,36,113} proposed that hydrogen bonds to O4a support deactivation via db isomerisation whereas hydrogen bonds to O1 support deactivation via sb rotation. These correlations are caused by different electronic structure of db and sb twisted minima and have been explained in section 4.3 of this chapter.

While the comparison of the average number of hydrogen bonds at the moment of excitation is promising, the number of hydrogen bonds for each individual simulation shows no clear trend as can be seen from tables 4.5 and 4.6.

For neutral Arg52, deactivations via db isomerisation are repressed compared to protonated Arg52. This corresponds to the observation that more hydrogen bonds to O1 are found in simulations with neutral Arg52 because water molecules are able to enter the chromophore binding pocket after the movement of the Arg52 side chain. These water molecules form hydrogen bonds with the easily accessible O1. O4a can not be approached by water

molecules when it is located in its X-ray structure position because of the steric hindrance of the amino acids surrounding it. The only exception is the water molecule bridging Thr50 to O4a. The resulting preference of sb rotation as deactivation event would result in lower productivity of the protein because no *cis*-conformer can be formed via this pathway. The missing positive charge might also be a reason for the preference of deactivation via sb rotation. In wild type PYP, the Arg52 side chain is positioned next to the chromophore phenyl ring. The positive charge of the side chain might stabilise the negative charge on the ring, which would favour the db twisted intermediate over the sb twisted intermediate.

The finding that the average hydrogen bond distance between O1 and Cys69 N is shorter for simulations where db isomerisation is observed than for simulations yielding deactivation via sb rotation is associated with the above described ability of water molecules from the bulk to form hydrogen bonds to O1. It results from the fact that more space is needed when more than one hydrogen bond is formed to O1 because a water molecule from the bulk entered the chromophore binding pocket and hydrogen bonds to O1.

As the differences between structural characteristics at the time of excitation are subtle more data is needed to varify the before specified tendencies whether simulations result in db isomerisation or sb rotation.

Deactivations via db isomerisation continue on the ground state either to *trans*- or to *cis*-conformation. Four out of ten simulations with db isomerisation adopt the *cis*-conformation for protonated Arg52 while it was one out of eight simulations for neutral Arg52. The remaining simulations relax to *trans*-conformation.

The ethylene chain is strained for the deactivation product in *cis*-conformation since the hydrogen bond between O1 and Cys69 N is not broken. The observation of the 90 degree twist between the C1 - O1 bond and the phenyl ring agrees with studies of Schotte *et al.*¹¹⁴. They managed to follow the structural changes in PYP using time-resolved Laue-crystallography while the protein completed the photocycle. With this method, they were able to detect intermediates during the early stages of the PYP photocycle. For the first intermediate detected 100 ps after excitation a twisted ethylene chain could be observed. The experimental results display that the twist is distributed evenly over all three participating bonds as has been observed in this work for three of five simulations which show the *cis*-product. In the remaining two structures, the double bond is planar and the twist is distributed over the two adjacent bonds.

Jung *et al.*¹¹⁵ independently studied the early parts of the PYP photocycle with the same method and observed a twisted intermediate after less than

100 ps as well. For comparison of the structures observed in the simulations described in this thesis and the twisted early intermediates of the PYP photocycle observed by Schotte *et al.* and Jung *et al.* two snapshots from simulations and the published crystal structures are depicted in Figure 4.23. For the two simulations, which show the strained *cis*-conformation with intact hydrogen bond to Cys69 N, this conformation is stable throughout the 10 ps ground state QMMM MD simulation. This simulation time is a factor of ten shorter than the 100 ps after which the experimental data has been collected. Continuation of these simulations could confirm the agreement of the theoretical observations in this thesis with the experimental results of Schotte and Jung *et al.*

Both experimental studies agree with our structural findings in three out of five simulations. The remaining two simulations differ from experimental data but not significantly. They show a strained *cis*-conformation with an intact hydrogen bond between O1 and Cys69. The fact that the twist is not distributed over the double bond could be due to the short simulation time during which this conformation was observed. More observed deactivations to the *cis*-conformation would yield better statistics on the probability distribution of the two twisted structures.

The subsequent simulations show three possible developments: first, re-isomerisation to *trans*-conformation, second, stable *cis*-conformation with broken hydrogen bond to Cys69 and third, stable *cis*-conformation with intact hydrogen bond to Cys69. It was proposed by van Wilderen *et al.*¹¹⁶ that breakage of the hydrogen bond between O1 of the chromophore and the Cys69 backbone is needed for the continuation of the photocycle. They performed femto-second visible pump/mid-IR probe spectroscopy on PYP in water. Thereby, they detected the appearance of a band after formation of the *cis*-isomer on ground state. This band is assigned to O1 without hydrogen bonds hence the hydrogen bond between O1 and the Cys69 backbone N^{117,118} has been broken. Intermediates for which this hydrogen bond did not break re-isomerised to the ground state. This intermediate state was termed GSI and can be assigned to deactivations via db isomerisation for which the hydrogen bond between the chromophore and Cys69 does not break. The majority of all simulations (69 %) with protonated Arg52 show a db twisted intermediate and would deactivate via db isomerisation on a longer timescale. The amount of deactivations via sb rotation is too small to account for the experimentally observed yield for the GSI intermediate of 0.58. From our calculations, it can be seen that re-isomerisation to the *trans*-configuration from a *cis*-configured intermediate is possible and does include an intact hydrogen bond to Cys69.

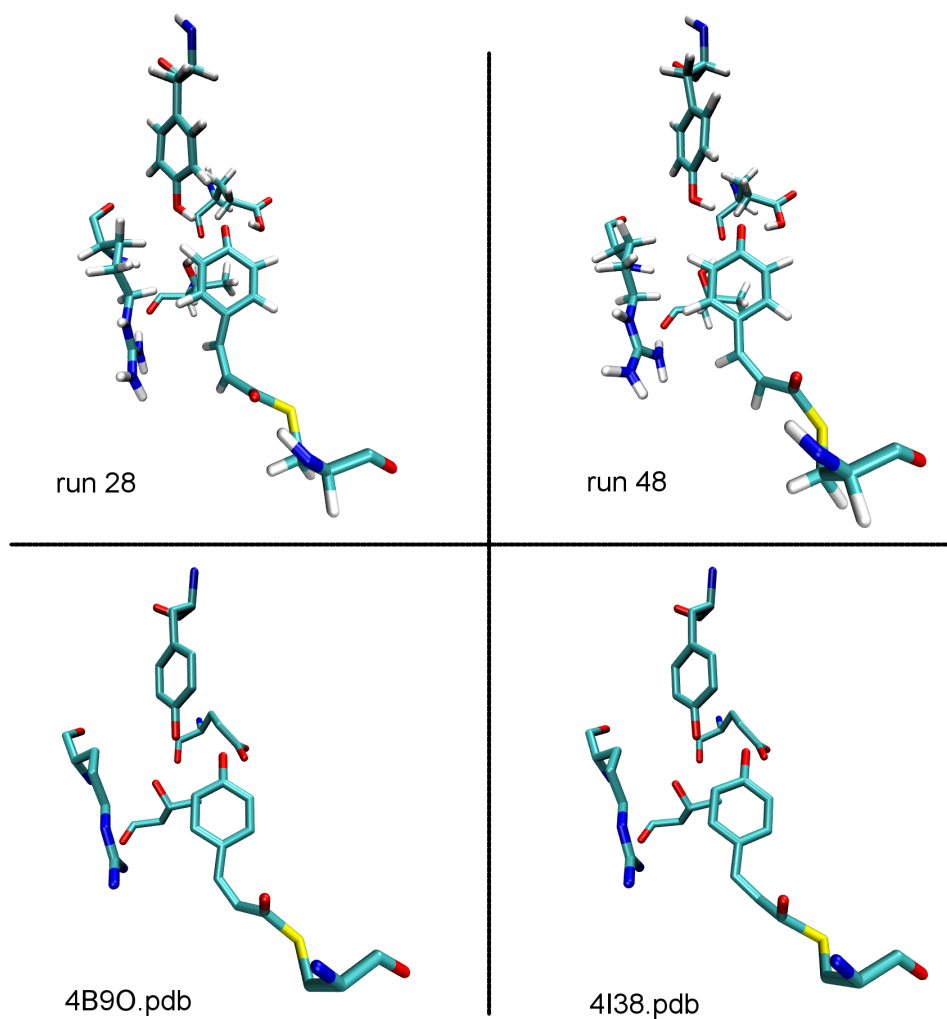


Figure 4.23: In excited QMMM simulations twisted intermediates were observed after deactivation via db isomerisation and subsequent rotation to the *cis*-conformation. Similar conformations were reported recently by Schotte *et al.* and Jung *et al.* through time-resolved Laue-crystallography. The theoretical results presented in this thesis match the conformation shown by the experimental studies.

The observations from the simulations described in this thesis are consistent because they show that it is possible to observe deactivation via db isomerisation to the *cis*-conformation and subsequent breakage of the hydrogen bond between O1 and the Cys69 N. Thus, the mechanisms of the first steps into a successful photocycle are accessible by theoretical methods.

When comparing the simulations, which showed deactivation via db isomerisation for protonated Arg52, to the study of Groenhof *et al.*²⁵ both setups show the same trends. The quantum yield of successful entry into the PYP photocycle, deactivation via db isomerisation ending in *cis*-conformation on the ground state, is 0.4 for the results presented in this thesis and 0.3 for the preceding work of Groenhof *et al.* These ratios are consistent with experimental studies where quantum yields of 0.35 to 0.5 were observed^{40,119}. The different level of theory used by Groenhof *et al.* and in this thesis makes a comparison arguable. However, as described in chapter 2.5 CAS 6,6 as used by Groenhof *et al.* is very sensitive to the choice of orbitals and the used basis set, 3-21G, is smaller than the one used in this thesis. Therefore, it is possible that the chosen set of active orbitals has biased the system in direction of deactivation via db isomerisation. Additionally, the Arg52 side chain position is different for the simulations presented in this thesis and the one observed by Groenhof *et al.* In their study they found the positively charged Arg52 side chain in a stacked conformation with the chromophore phenyl ring. This may be an artefact of the used GROMOS43a2 force field. Personal communication revealed that the stacked conformation was not observed for recent simulations of Groenhof with the AMBER03 force field. The simulations in this thesis with the Amber03 force field show the Arg52 side chain next to the chromophore with the planes spanned by the phenyl ring and Arg52 side chain perpendicular to each other. The stacked conformation results in a lower distance between the chromophore and Arg52 which is likely to influence the coulomb interactions between these two charged residues. These interactions might stabilise the negative charge on the phenyl ring in the excited state which would favour deactivation via db isomerisation.

In this thesis, the fluorescence decay time was calculated by fitting the function $f(t) = \exp(-t/\tau)$ to the time it took the protein to decay via db isomerisation. The time used for this calculation was the moment of the surface hop from excited to ground state. The fluorescence decay time calculated was 1377 fs. Experimentally, it was measured by femtosecond fluorescence up-conversion measurements¹²⁰. The fastest component was 430 fs while the second component was 1 to 3 ps, rising with lower temperature. The calculated fluorescence decay time is in the regime of the experimental measurements. The used surface hopping algorithm might lead to a larger value because it

underestimates the crossing probability.

However, deactivations via sb rotation observed in this study have been neglected in this comparison. Taking them into account, the quantum yield for protonated Arg52 is diminished to 0.15. This result may be an indication that the setup used in this work is not sufficient for simulating the deactivation events in PYP. On the other hand, the time scale used might be too short because in 48 of the 61 simulations, which did not show deactivation during the simulation time, the chromophore is in the db twisted conformation. Deactivation of all simulations could give a higher quantum yield again depending on the ratio between *cis*- and *trans*-conformation. Longer simulation times would be the first step to look into this issue, the next step being variations of the QM method.

Rigid scan calculations on the chromophore model in vacuum present an effect of the hydrogen bond length on the heights of the barriers in the excited state between the planar Franck-Condon-region and the two twisted minima. The remarkable short length observed by X-ray crystallography favours deactivation via db isomerisation compared to longer hydrogen bonds by reducing the barrier along this deactivation pathway. The contrary is observed for sb rotation, where the barrier is lower in energy when the hydrogen bond length is in range of typical hydrogen bonds as described in chapter 1.2. Thus, for longer hydrogen bonds, deactivation via sb rotation might be enhanced because the long hydrogen bonds do not support deactivation via db isomerisation efficiently.

Gromov *et al.*¹¹³ calculated the barriers between the Franck-Condon-region and the twisted intermediates by computing the minimum energy pathway along the db and sb torsion. The hydrogen bonds from the chromophore to Tyr42 and Glu46 were 2.604 and 2.600 Å long, respectively. This length is in the range observed experimentally by Yamaguchi *et al.*. They compared the barriers computed from this setup to barriers where the two mentioned hydrogen bonds were formed between the chromophore and two water molecules. In this case, the hydrogen bonds were longer, 2.725 and 2.724 Å, respectively. By going from the two water molecules as hydrogen bonding partners to Tyr42 and Glu46, the barrier for db torsion decreased while the barrier for sb torsion was unchanged. The barrier for db torsion was still higher than for sb torsion. These results lead to the conclusion that shorter hydrogen bonds stabilise the pathway of db torsion more than the sb torsion pathway which is consistent with the results from the rigid scans.

As the calculations were carried out using a model system and the rigid scan method, the observations described above show the tendency for shortening of the hydrogen bond length but no quantitative values. The barrier heights

in the relaxed protein will differ. These calculations draw interest to the outcome of excited state QMMM MD simulations for a system with bond lengths in the regime of the hydrogen bonds published by Yamaguchi *et al.*

The in this chapter mentioned rates for the deactivation events could change significantly for neutral Arg52 as well when not deactivated simulations, which show twisted conformations, would deactivate on a longer time scale. It can be speculated due to the observed twisted intermediates that the effect of the Arg52 deprotonation on the deactivation pathway and thus completion of the photocycle is larger than already shown. Continuation of the simulations might give new insights and better statistics on this matter, too.

The findings of this chapter are in agreement with the common belief that Arg52 in PYP is protonated. For protonated Arg52 the protein in water is stable throughout our simulations whereas large conformational changes from the X-ray structure are observed for neutral Arg52. Additionally, the deactivation events from the excited to the ground state are altered from primarily db isomerisation to ab rotation after Arg52 deprotonation. Whether PYP can be stable with neutral Arg52 under the crystallisation conditions of Yamaguchi *et al.* can not be verified by our setup and is subject of further studies.

The effect of the environment on a model system of the PYP chromophore will be described in the following chapter.

Chapter 5

pCK⁻ Deactivation in Water and Decanol

Several model systems of the PYP chromophore in solution have been studied by experimental and theoretical methods before. These studies help to understand the events during the photocycle. The advantage of model systems is that small changes like a different strength of the hydrogen bonds to the chromophore can be obtained without worrying about the influence on other parts of the protein. However, these studies only show trends which might hold in the protein because of the missing protein environment and its effects on the chromophore. In the following, the results of these previous studies are briefly summarised.

Espagne *et al.*^{32,121} studied several chromophore models in different solvents with Steady-State spectroscopy and Transient Absorption Spectroscopy. The chromophore models differed from each other in the electron acceptor strength of the carbonyl substituent. As solvents, they used water, ethylen glycol, DMF, and primary alcohols ranging from methanol to 1-decanol. They studied differential absorbance spectra for the different chromophore models which they attributed to two possible deactivation pathways. These observations were independent from the solvent. For chromophore models with a carbonyl substituent with a low electron acceptor strength deactivation to the *cis*-conformer was observed without detection of an intermediate. With a carbonyl substituent with high electron acceptor strength no *cis*-conformer was formed instead the formation of a transient state was observed. Based on these experiments, Espagne *et al.* proposed two relaxation pathways: 1) for weak electron acceptor substituents a concerted motion of the double bond and the adjacent bond in direction of the carbonyl group and 2) for strong electron acceptor substituents only torsion around the double bond.

Deactivation of the chromophore model pCK⁻ in water by QMMM MD sim-

ulations of Boggio-Pasqua *et al.*²⁷ revealed predominantly deactivation via sb rotation in 80 simulations whereas 11 simulations deactivated via db isomerisation. Closer inspection of the trajectories revealed that the hydrogen bond structure around pCK⁻ differs for the two deactivation pathways. For deactivation via sb rotation, three hydrogen bonds to O1 and one hydrogen bond to O4a were observed. For deactivation via db isomerisation two hydrogen bonds were formed by both O1 and O4a. The hydrogen bonds are proposed to stabilise the two twisted states on the excited state.

Experimental studies by Changenet-Barret *et al.*¹²² proposed different deactivation routes for pCK⁻, deactivation via sb rotation in water and via db isomerisation in decanol. The spectra for the photoreaction in both solvents were similar but different kinetics were observed for the occurring reactions. For water as solvent, the deactivation was interpreted to occur via a short-lived phenolate-twisted hot ground state because no long-lived intermediate, which would be the *cis*-conformer, is formed. This corresponds to deactivation via sb rotation. In decanol, a long-lived intermediate was detected. This isomer was interpreted as the *cis*-isomer formed upon deactivation via db isomerisation.

For the calculations performed for this thesis, the effect of the environment on the chromophore deactivation from the excited state was studied for the chromophore model pCK⁻, which was also used in chapter 2.5 (Figure 2.3), in water and decanol. This approach was chosen because these solvents have the largest differences in their chemical properties of the solvents used in experimental studies on chromophore models. Furthermore, the detailed study on these systems of Changenet-Barret *et al.* enables the comparison to experimental data.

5.1 pCK⁻ in Water

To study the photoreaction pathway for pCK⁻ molecules in water, the molecule together with a sodium ion was soluted in 4087 water molecules in a rectangular box with 5 nm edge length. For this system a ground state MM MD simulation of 50 ns in steps of 0.002 ps was carried out using the GROMOS43a2 force field and the SPC water model. Each 0.5 ns the structure was extracted and used as input structure for subsequent ground state QMMM MD simulations. These calculations were simulated for 1000 steps of 0.0005 ps with CAS 6,6/3-21G as the QM method using the GROMACS/GAUSSIAN 03 interface. The MM part was described by the GROMOS43a2 force field with

SPC water model.

From this all successfully finished runs were taken, 68 in total, and the output geometry was used as input structure for subsequent runs. At first an equilibration with 5000 steps of 0.001 ps was carried out with RHF and a 3-21G basis set. Followed by 5000 steps of 0.001 ps with RHF/6-31G*. The production runs starting on the excited state were carried out for steps of 0.0005 ps with CAS 8,8 and a basis set of 6-31G* for the QM region using the GROMACS/MOLPRO interface. The MM part was described by the same force field as the previous ground state calculations.

For the non-bonded interactions, van-der-Waals and Coulomb interactions, cut-offs were used. All interactions between atoms closer than these cut-offs were calculated at each step. Beyond this cut-off coulomb interactions were treated with a reaction-field correction with an infinite relative dielectric constant. For all simulations except the excited state QMMM MD production runs a cut-off of 1.6 nm was used for both interactions while a cut-off of 2.0 nm was applied in the production runs.

The orbitals for CASSCF were the same for pCK⁻ in water and decanol.

The results from the MD trajectories are summarised in Table 5.1. The chromophore was found to undergo both deactivation pathways, db isomerisation and sb rotation. Deactivation via sb rotation was observed for 39 simulations while the chromophore deactivated by db isomerisation in 16 simulations. Most of these deactivation events relaxed to the initial planar *trans*-conformation after crossing from excited state to gs, about 63 % for db isomerisation and 76 % for sb rotation. The remaining 6 simulations for db isomerisation and 8 for sb rotation end in *cis*-conformation and a 180 degree twist of the phenyl ring, respectively.

For 13 simulations no twists around either sb or db larger than 50 degrees could be observed. This simulations did not cross from excited state to ground state during simulations times between 1.4 and 3.0 ps and retained a planar conformation on the excited state.

Deactivations from the excited state via db isomerisation took place immediately after the chromophore conformation had reached the db twisted conformation on the excited state. For deactivation via sb rotation the time between reaching the sb twisted minimum and the surface hop reached from no interim time to 1.250 ps.

The average time until deactivation from the excited state took place is longer for sb rotation than for db isomerisation, 0.773 ps and 0.356 ps respectively. Before the twist occurred that lead to successful deactivation from the excited state no twisting movement around another bond than the one that finally lead to deactivation could be observed in any trajectory.

Table 5.1: Summary of the simulation results for the chromophore model pCK⁻ in water. No twists could be observed for 12 simulations, which corresponds to 17.65 %.

deactivation event		db	sb	total
count		16	37	68
%		23.53	54.41	-
to planar gs (count)		10	27	37
to planar gs (%)	of all runs	14.71	39.71	54.41
	of deactivation event	62.50	72.97	69.81
to cis/180 degrees (count)		6	8	14
to cis/180 degree (%)	of all runs	8.82	11.76	20.59
	of deactivation event	37.50	21.62	26.42
average time till sh (steps)		715.63	1546.38	-
average time till sh (ps)		0.356	0.773	-

At the moment of excitation up to five hydrogen bonds to O4a and four hydrogen bonds to O1 were counted using the `g_hbond` utility implemented in GROMACS. The criteria of `g_hbond` are a maximum of 3.5 Å for the donor - acceptor distance and 30 degree for the angle A-B-H. Note that the number of hydrogen bonds found according to these criteria is unusually high. For simulations resulting in deactivation via db isomerisation on average 4.00 hydrogen bonds to O4a and 2.44 hydrogen bonds to O1 were found. Simulations showing deactivation via sb rotation had 3.89 hydrogen bonds to O4a and 3.00 hydrogen bonds to O1.

The results are in line with the observations of deactivations of the chromophore in the protein presented in section 4.3.

Table 5.2: Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for pCK⁻ in water.

deactivation event	average number of HBs to	
	O1	O4a
db	2.44	4.00
sb	3.00	3.89

5.2 pCK⁻ in Decanol

For comparison to the experimental data of Chaugenet-Barret *et al.*¹²², simulations and analysis for pCK⁻ in decanol, performed by Lela Vukovic¹, is described in the following section.

For simulations in decanol, pCK⁻ was placed in a box with 564 decanol molecules. As force field for all the performed simulations the same force field as for the simulations of pCK⁻ in water, i.e. the GROMOS53a6 force field, was used. The box, with an edge length of 7.5 nm, was equilibrated for 20 ns. Subsequently a 100 ns MM MD simulation was started. Every nanosecond coordinates and velocities were extracted and used as input structure for the following QMMM MD simulations.

The 99 obtained input structures were applied to start ground state QMMM MD simulations. For each structure a simulation of 3000 steps of 0.0005 ps was carried out using RHF/6-31G* as QM method for the chromophore model. Successively excited state QMMM MD simulations were started from the ground state QMMM MD simulations with a step size of 0.0005 ps. The chromophore model was described by CAS 8,8/6-31G* as QM method.

For pCK⁻ in decanol the observed behaviour differs from that of pCK⁻ in water. During an average simulation time of 2.16 ps no deactivation via sb rotation occurred. 42 simulations (43.75 %) showed twisting around the sb of more than 75 degrees but did not cross from excited state to ground state. Deactivation via db isomerisation was observed in 21 simulations (21.875 %) and the hops from excited state to ground state took place immediately after pCK⁻ reached the twisted conformation. For 33 simulations (34.375 %) no twists larger than 50 degrees around db or sb were observed. The described information is summarised in Table 5.3.

Simulations of pCK⁻ in decanol showed less hydrogen bonds to O1 at the moment of excitation when deactivation via db isomerisation was observed compared to simulations which showed a twist around the sb. The contrary was observed for hydrogen bonds to O4a. (Table 5.4)

¹University of Illinois at Urbana-Champaign

Table 5.3: Summary of the simulation results for the chromophore model pCK⁻ in decanol. No twists could be observed for 33 simulations, which corresponds to 34.375 %.

deactivation event		db	sb	total
count		21	42	96
%		21.875	43.75	
to planar gs (count)		11	-	-
to planar gs (%)	of all runs	11.46	-	-
	of deactivation event	52.38	-	-
to cis/180 degrees (count)		10	-	-
to cis/180 degree (%)	of all runs	10.42	-	-
	of deactivation event	47.62	-	-

Table 5.4: Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for pCK⁻ in decanol.

deactivation event	average number of HBs to	
	O1	O4a
db	1.00	2.67
sb	1.12	2.50

5.3 Discussion

The performed QMMM MD simulations of the excited chromophore model pCK⁻ in water and decanol revealed different pathways for the deactivation from the excited state. While pCK⁻ in water deactivated via both db isomerisation and sb rotation, in decanol only deactivation via db isomerisation was observed. The db or sb twisted intermediates on the excited state were observed for both solvents. These observations lead to the conclusion that twists around both db and sb are possible in water and decanol but the deactivation from the sb twisted minimum is inaccessible in decanol during the simulation time. The reason is the different charge localisation on the chromophore in the planar, db twisted, and sb twisted intermediates. In both the planar and the db twisted intermediates the negative charge is localised on the phenyl ring while it is located on the ethylene chain in the sb twisted intermediate. This charge translocation has been described in section 4.3 of this thesis. The solvent molecules have to arrange according to the charge localisation on the chromophore. As the charge localisation of the planar and the db twisted intermediate is similar, the required movement to adjust to changes is less than for the sb twisted intermediate which leads to an immediate deactivation from the excited state after formation of the db twisted intermediate. Water molecules are smaller than decanol molecules and the viscosity of the solvent is lower which leads to a faster rearrangement upon the charge translocation from the sb twist. The duration of the simulations presented in this thesis might be too short for the decanol molecules to rearrange to enable deactivation from the sb twisted intermediate.

For both solvents the sb twisted intermediate is observed twice as often as the db twisted intermediate. Most deactivations, 70 % in water and 52 % in decanol, end in *trans*-conformation on the ground state. The value for decanol is lower because only deactivations via db isomerisation were observed and included. 8.82 % of all simulations for water and 10.42 % for decanol showed deactivation via db isomerisation with subsequent formation of the *cis*-conformation on the ground state.

These values are similar enough for us to conclude that the deactivation mechanisms for the chromophore model pCK⁻ in water and decanol are the same. The different observations result from the differing viscosity of the solvents which slows down the deactivation via sb rotation in case of decanol.

QMMM MD simulations of pCK⁻ in water with a similar setup by Boggio-Pasqua *et al.*^{27,123} using CAS 6,6/3-21G for the chromophore and the SPCE force field¹²⁴ for the water molecules yielded different ratios for deactivation via db isomerisation and sb rotation. Of 91 simulations 11 deactivated via db

isomerisation and 80 via sb rotation, which are 12.1 % and 87.9 %, respectively. In this study all simulations showed deactivation from the excited state. Seven simulations where the chromophore deactivated via db isomerisation showed subsequent evolution to the *cis*-conformation while four evolved back to *cis*-conformation on the ground state.

A possible explanation for this discrepancy is that the potentials on the excited state surface are steeper when using CAS 6,6 than when using CAS 8,8. This might lead to an easier accessibility of both twisted intermediates. As a result of this higher accessibility all simulations reach a conical intersection during the simulation time and show deactivation from the excited state. The differing ratios of deactivations via sb rotation and db isomerisation while applying CAS 6,6 compared to CAS 8,8 could also be due to different barrier heights for these two active space sizes. The effect of different active space sizes on the barriers on the excited state potential energy surface is described in chapter 2.5 of this thesis.

The aforementioned experimental studies on the chromophore model pCK⁻ in water and decanol have been published by Changenet-Barret *et al.*¹²². They studied the system with time-resolved transient absorption spectroscopy, by measuring the development of the excited state absorption (ESA) and stimulated emission (SE) for up to 1 ns for wave lengths between 300 and 700 nm.

In the experiment, pCK⁻ was solvated in a basic aqueous solution with pH 10.1. For this setup the following was observed. The band of the excited state absorption was centred at 350 nm and the band of the stimulated emission at 480 nm both bands decay for the whole observation time of 50 ps. Simultaneously the ESA band shifts to the blue while the SE band shifts to the red. Up until a measuring time of 2.5 ps a transient absorption band rises around 430 nm. This band subsequently decays between 2.5 and 50 ps. For pCK⁻ in decanol the overall picture resembles the observations of pCK⁻ in water differing in the time scales and thus support the findings presented here. Both ESA and SE bands decay while shifting towards the blue and the red, respectively. This decay takes place during the first 500 ps of measurement while the bands show no further reduction until the end of the experiment after 1 ns. The transient absorption band rises during the first 75 ps and decays afterwards. The maxima of these three bands are shifted compared to the measurement of pCK⁻ in water: the ESA band is centred at 360 nm, the SE band at 460 nm and the transient absorption band at 440 nm.

They interpreted their results as follows. In water the transient absorption band at 430 nm is assigned to the phenolate-twisted hot ground state. A

deactivation via db isomerisation is excluded based on the observed sensitivity to the polarity and viscosity of the solvent of the analogous band in the spectrum of the chromophore model pCT⁻³² where no formation of a *cis*-conformer could be observed. The spectra of pCT⁻ and pCK⁻ in water show great similarity which leads to the suggestion by Changuenet-Barret *et al.* that the mechanisms for both chromophore models are analogue. For pCK⁻ in decanol the long-lived state showing in the spectra as transient absorption band at 440 nm is assigned to the formation of a *cis*-conformer. Decanol as solvent was proposed to hinder deactivation from the excited state via sb rotation sterically leading to solely deactivation via db isomerisation. The *cis*-conformer is not stable and re-isomerises on the ground state to the *trans*-conformer. The barrier for the re-isomerisation was estimated to have a height of about 84 kJ/mol. The different mechanism for the deactivation from the excited state for pCK⁻ in water and decanol are assigned to four differing physical and chemical properties of the solvents: polarity, solvation dynamics, viscosity and hydrogen bond donor character. Our simulations suggest that the viscosity and the subsequent solvation dynamics are the reason behind the differing observations for the deactivation of pCK⁻ in water and decanol. The polarity and from this the hydrogen bond donor character seem to be of minor importance because they would influence the ration between observed db and sb twisted intermediates which is the same for both solvents.

Dmitry Morozov¹²⁵ calculated the absorption, ESA, and SE bands for pCK⁻ solvated in EFP water¹²⁶. The SE band for *trans*-configured chromophore is found at 490 nm, corresponding to the experimental SE band at 480 nm. The experimental ESA band could not be reproduced because the calculations took into account only the first four excited states. Extrapolating from data obtained for these states, the ESA band observed in experiment would correspond to a excitation from the first to the about eighths excited state. The *cis*-conformer is excited from the ground state at 405 nm while the SE band for the sb-twisted state is found at 680 nm. For pCK⁻ in decanol, absorptions at higher wave numbers are expected. Further calculations on this matter are work in progress.

In summary, merging the experimental data with the theoretical results leads to the model that both pCK⁻ in water and in decanol shows deactivation via db isomerisation. The transient absorption band at 440 and 430 nm, respectively, originates from the *cis*-conformer. The absorption bands of the sb-twisted state are found at wave length larger than 700 nm and are therefore not measured in the experiments. The different time scales for the

formation of the *cis*-conformer are due to the different viscosity of the solvents. The fast extinction of the transient absorption band in water is object to further research. Calculations for the chromophore isomerisation from *cis*-to *trans*-conformation in water showed a longer time scale of several ns. A possible explanation would be the participation of hydroxid ions or protons which would catalyse the process.

To resolve whether our interpretation of a similar mechanism for the pCK⁻ deactivation in both solvents or the interpretation by Changenet-Barret *et al.* are accurate experimental studies which include wavelengths above 700 nm and theoretical calculations with longer simulation times are necessary.

Chapter 6

Conclusion & Outlook

Several ways how the protein environment in PYP influences the hydrogen bond network around the chromophore and the deactivation of the chromophore from excited state to ground state were investigated in this thesis. These investigations were all based on theoretical methods, QM calculations and MD simulations, on the protein and on model systems. In particular, this three main questions were addressed:

- 1) Can the experimentally observed SSHB be reproduced by theoretical methods and which conditions are required?
- 2) How does the protonation state of Arg52 influence the protein dynamics and deactivation pathway?
- 3) How does the hydrogen bonded network around the chromophore affect the deactivation from the first excited state?

In the following, the findings of this thesis regarding these questions are summarised:

1) Can the experimentally observed SSHB be reproduced by theoretical methods and which conditions are required?

Chapter 3 explored the hydrogen bond between the chromophore and Glu46 by calculating the potential energy of the hydrogen position along the bond. We found that the Arg52 protonation state modulates the hydrogen bond. Compared to the modulation by the Arg52 protonation state, other variations in the computational setup have small effect on the resulting potential energy profile. The general trend observed was that for neutral Arg52 the proton is delocalised over two positions, bound to the chromophore and bound to Glu46. Protonated Arg52 causes the proton to bind to Glu46. The results were compared to experimental studies of Yamaguchi *et al.* and theoretical works of Saito *et al.*, which disagreed on the Arg52 protonation state and

the position of the proton in the mentioned hydrogen bond. Our results were in agreement with both contradicting studies because the Arg52 protonation state showed the most distinct effect of all parameters on the energy potential and is correlated to the proton position. However, it is not clear yet why Arg52 is found to be deprotonated by Yamaguchi *et al.*. The energy gain of the delocalisation of the proton, 3 kJ/mol, is too small to explain the deprotonation. Theoretical calculations on the effect of the crystallisation on the pK_a of Arg52 by Matthias Ullmann (personal communications) is half a pK_a unit. Further work is needed to address this important issue.

2) How does the protonation state of Arg52 influence the protein dynamics and deactivation pathway?

The strong influence of the Arg52 protonation state on the potential energy curve along the hydrogen bond between the chromophore and Glu46 and the experimental observation that Arg52 is neutral under the conditions of Yamaguchi *et al.* led to the question how the protonation state of Arg52 influences the protein dynamics and deactivation events in water. These processes were investigated in chapter 4 by using MM and QMMM MD simulations. Simulations in solution showed that deprotonated Arg52 leads to a less stable chromophore binding pocket. The hydrogen bond between Arg52 and Tyr98 was broken and the chromophore binding pocket opened to the bulk water similar to X-ray structures of the signalling state pB. As a result, bulk water molecules were able to enter the binding pocket and formed hydrogen bonds to O1. In five of six MM MD simulations the direct hydrogen bond between Thr50 and O4a, which was also observed in MM MD simulations with protonated Arg52, was broken and reformed as indirect hydrogen bond bridged by a bulk water molecule, which was stable at this position throughout the simulations. This hydrogen bond between the chromophore and Thr50 has not been observed by X-ray crystallography. It is probably formed to stabilise the chromophore charge because the hydrogen bonds to Tyr42 and Glu46 are longer in the simulations than in X-ray structures. Force fields parametrised for shorter hydrogen bond length at these positions or usage of other QM methods might reproduce the experimentally observed hydrogen bond length and as a result it can be reviewed whether the hydrogen bond to Thr50 exists also with short hydrogen bonds to Tyr42 and Glu46. For protonated Arg52, the chromophore binding pocket was stable in all MM MD simulations. These observations lead us to the conclusion that the protein structure with neutral Arg52 is violated in solution.

Further, the effect of the Arg52 deprotonation on the deactivation events from

the first excited state was investigated. For protonated Arg52 primarily deactivation via db isomerisation was observed whereas neutral Arg52 favoured deactivation via sb rotation. The majority of all simulations did not show deactivation from the excited state during the simulation time. The observed twisted structures in these simulations indicate that longer simulation times might show an even more distinct picture of the Arg52 protonation state influence on the deactivation events.

For both protonation states, the occurring deactivation event showed correlations to the number of hydrogen bonds formed by O4a and O1 of the chromophore at the moment of excitation. On average more hydrogen bonds to O4a favour db isomerisation whereas more hydrogen bonds to O1 favour the sb rotation pathway. Water molecules, which entered the chromophore binding pocket for neutral Arg52, formed more hydrogen bonds to O1 and, as a result, sb rotation was preferred for neutral Arg52. Thus the Arg52 protonation state has a clear influence on the deactivation but it is not obvious whether this is due to the different charge or exposing of the chromophore to the bulk water.

40 % of all observed deactivations via db isomerisation continued to a strained *cis*-conformation on ground state with intact hydrogen bond between O1 and Cys69 N. This conformation corresponds to recent experimental results. Continuation of these simulations on ground state showed breakage of the hydrogen bond in two cases which resulted in a planar *cis*-conformation. One simulation returned to *trans*-conformation while the remaining two maintained the strained *cis*-conformation.

The results for water-soluted PYP in both ground and excited state show that the deprotonation of Arg52 destabilises the chromophore binding pocket already on ground state. Such a destabilisation with its subsequent conformational changes has not been observed experimentally during the photocycle before formation of the signalling state pB^{10,21,104}. In contrast to simulations with protonated Arg52, deactivation from the excited state for neutral Arg52 takes place mainly via sb rotation instead of db isomerisation which would lead to successful entry into the photocycle. The conformation of the hydrogen bond network surrounding the chromophore at the moment of excitation controls which deactivation pathway takes place. The observed mechanics for deactivation via db isomerisation with a strained *cis*-conformation, successive breaking of the hydrogen bond between O1 and Cys69 N and continued rotation to planar *cis*-conformation agrees with recent experimental studies^{114,115}.

The conclusion of the described results is that the deprotonation of Arg52 destabilises PYP in water and changes its deactivation processes in a way that the pathway via db isomerisation is repressed. From our simulations of

PYP in water, Arg52 has to be protonated to remain stable and behave in its natural manner.

3) How does the hydrogen bonded network around the chromophore affect the deactivation from the first excited state?

The impact of the hydrogen bonded network on the chromophore deactivation from the excited state was studied with QMMM MD simulations of the chromophore model pCK⁻ in water and decanol, presented in chapter 5. In both solvents, deactivation from the excited state via db isomerisation has been observed. After the db twisted conformation on the excited state was reached, deactivation to the ground state took place immediately. Deactivation via sb rotation has only been observed for pCK⁻ in water. The sb twisted conformation was stable on the excited state for up to 1.25 ps before deactivating to the ground state. In decanol the sb twisted conformation was formed but no deactivation took place during the QMMM MD simulations. Deactivations via db isomerisation ended in *cis*-conformation on ground state in 37.5 % of the simulations in water and 47.6 % for simulations in decanol. This ratios are similar to the ratio for PYP in water presented in this thesis. Another similar observation are the average numbers of hydrogen bonds at the moment of excitation. For both pCK⁻ in water and in decanol more hydrogen bonds to O4a have been found for simulations with db twisted intermediates than for simulations with sb twisted intermediates. The opposite was found for the hydrogen bonds to O1. These observations correspond to the results of the excited state QMMM MD simulations of PYP in water, described above.

The ratio of db versus sb twisted geometries observed on the excited state is in the same range for pCK⁻ in water, 0.43, and decanol, 0.5. Combining this with the fact that the average numbers of hydrogen bonds at the moment of excitation show the same pattern for both solvents indicates that the hydrogen bonded network around the chromophore triggers which twist occurs and not the solvent.

From our calculations we can conclude that for both solvents deactivation via db isomerisation takes place. Deactivation via sb rotation could only be observed for pCK⁻ in water during the simulation time. This is due to the different viscosity of the two solvents. Water molecules are able to rearrange faster than decanol molecules and thus stabilise the, in comparison to ground state changed, charge distribution of the sb twisted state during the simulation time. This enables the deactivation via this pathway during our simulations. MM MD simulations showed that decanol molecules need about

100 ps to rearrange around sb twisted pCK⁻ which is too slow for QMMM MD simulations. Longer simulation times are required to observe whether deactivation via sb rotation is possible for pCK⁻ in decanol.

The comparison of our simulation results with the experiments shows that the deactivation mechanism described in this work fit to the experimental data although the experiments were interpreted in a different way. Experiments with a broader wavelength distribution could help to resolve this issue.

With the investigations of this thesis, a better understanding of the photocycle of PYP, exemplary for the PAS domain super family and for photoactive proteins, is possible. Future studies on the PYP photocycle should particularly pay attention to the interactions of the chromophore with its environment. Comprehension of the photoactivated mechanisms in PYP are important for understanding biological processes in other photoactive proteins and further members of the PAS domain super family. Applications like controlling bio-molecular events with light could become feasible.

Chapter 7

Appendix

List of Abbreviations

A	hydrogen donor
Arg52	Arginine 52
B	hydrogen acceptor
CASSCF	Complete Active Space Self-Consistent Field
CI	Configuration Interaction
CIS (D,T,Q)	CI Singles (Doubles, Triples,Quadruples)
Cys69	Cysteine 69
db	double bond
DFT	Density Functional Theory
DNA	Deoxyribonucleic Acid
EC	Electron Correlation
es	excited state
FTIR	Fourier Transform Infrared Spectroscopy
GGA	Generalised Gradient Approximation
Gln41	Glutamine 41
Glu46	Glutamic Acid 46
gs	ground state
H	hydrogen atom
hb	hydrogen bond
HF	Hartree Fock
HK	Hohenberg and Kohn
Ile58	Isoleucine 58
KC	Kohn and Sham
Lys123	Lysine 123

LBHB	Low Barrier Hydrogen Bond
LDA	Local Density Approach
MCSCF	Multi-Configuration Self-Consistent Field
MD	Molecular Dynamics
MM	Molecular Mechanics
MO	Molecular Orbital
PAS	Per-Arnt-Sim
pB state	signalling state of PYP
pCA ²⁻	chromophore model <i>p</i> -hydroxycinnamate
pCM ⁻	amide analogue of chromophore model <i>p</i> -hydroxycinnamate
PDB	Protein Data Bank
PES	Potential Energy Surface
pG state	ground state of PYP
Phe79	Phenylalanine 79
pR state	excited state of PYP
PYP	Photoactive Yellow Protein
QM	Quantum Mechanics
QMMM	Quantum Mechanics/Molecular Mechanics
RHF	restricted Hartree Fock
sb	single bond
SCF	Self-Consistent Field
SD	Steepest descend
Ser8	Serine 8
Ser117	Serine 117
SSHB	Short Strong Hydrogen Bond
Thr50	Threonine 50
Tyr42	Tyrosine 42
Tyr98	Tyrosine 98
Tyr118	Tyrosine 118

Influence of Arg52 protonation state on PYP dynamics and chromophore isomerisation

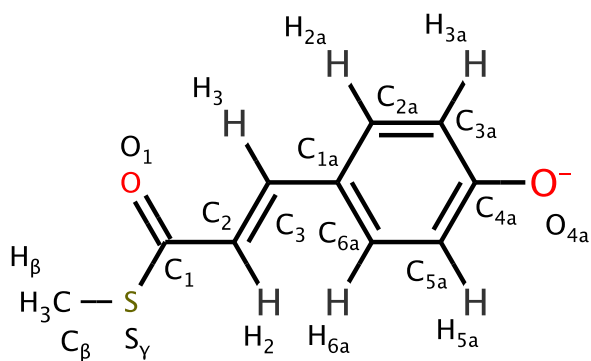


Figure 7.1: Depiction of the planar *trans*-configured chromophore with atom names.

Table 7.1: Amber atom type and charge for the PYP chromophore parametrised by MP2 geometry optimisation and subsequent determination of the charges by restricted electrostatic potential fitting (RESP).

atom name	amber atom type	charge
C1	amber99_2	0.719
O1	amber99_41	-0.589
C2	amber99_3	-0.642
H2	amber99_22	0.206
C3	amber99_3	0.130
H3	amber99_22	0.074
C1a	amber99_3	-0.078
C2a	amber99_3	-0.105
H2a	amber99_22	0.113
C3a	amber99_3	-0.425
H3a	amber99_22	0.114
C4a	amber99_2	0.777
O4a	amber99_45	-0.775
C5a	amber99_3	-0.448
H5a	amber99_22	0.121
C6a	amber99_3	-0.02
H6a	amber99_22	0.093

Table 7.2: Bond length between the atoms of the PYP chromophore for use in the AMBER03 force field calculated by MP2 optimisation

bond	length (Å)
$C_\beta - H_\beta$	1.09017
$C_\beta - S_\gamma$	1.80662
$S_\gamma - C_1$	0.83958
$C_1 - O_1$	1.23563
$C_1 - C_2$	1.43266
$C_2 - H_2$	1.08717
$C_2 - C_3$	1.37647
$C_3 - H_3$	1.09373
$C_3 - C_{1a}$	1.41591
$C_{1a} - C_{6a}$	1.42558
$C_{6a} - H_{6a}$	1.09097
$C_{6a} - C_{5a}$	1.37344
$C_{5a} - H_{5a}$	1.08938
$C_{5a} - C_{4a}$	1.45578
$C_{4a} - O_{4a}$	1.26100
$C_{4a} - C_{3a}$	1.45173
$C_{3a} - H_{3a}$	1.08911
$C_{3a} - C_{2a}$	1.37431
$C_{2a} - H_{2a}$	1.09263
$C_{2a} - C_{1a}$	1.42442

Table 7.3: Angles present in the PYP chromophore for use in the AMBER03 force field calculated by MP2 optimisation

angle	value (degree)
$H_{\beta} C_{\beta} S_{\gamma}$	110.551
$C_{\beta} S_{\gamma} C_1$	97.403
$S_{\gamma} C_1 O_1$	118.655
$S_{\gamma} C_1 C_2$	112.813
$O_1 C_1 C_2$	128.532
$C_1 C_2 H_2$	117.830
$C_1 C_2 C_3$	120.073
$H_2 C_2 C_3$	122.097
$C_2 C_3 H_3$	114.611
$C_2 C_3 C_{1a}$	129.396
$C_3 C_3 C_{1a}$	115.992
$C_3 C_{1a} C_{6a}$	123.985
$C_3 C_{1a} C_{2a}$	119.481
$C_{6a} C_{1a} C_{2a}$	116.535
$C_{1a} C_{6a} C_{5a}$	121.746
$C_{1a} C_{6a} H_{6a}$	119.088
$C_{5a} C_{6a} H_{6a}$	119.167
$C_{6a} C_{5a} H_{5a}$	120.747
$C_{6a} C_{5a} C_{4a}$	122.872
$H_{5a} C_{5a} C_{4a}$	116.380
$C_{5a} C_{4a} O_{4a}$	122.659
$C_{5a} C_{4a} C_{3a}$	114.133
$C_{3a} C_{4a} O_{4a}$	123.208
$C_{4a} C_{3a} H_{3a}$	116.741
$C_{4a} C_{3a} C_{2a}$	122.914
$H_{3a} C_{3a} C_{2a}$	120.945
$C_{3a} C_{2a} H_{2a}$	119.650
$C_{3a} C_{2a} C_{1a}$	122.400
$H_{2a} C_{2a} C_{1a}$	117.950

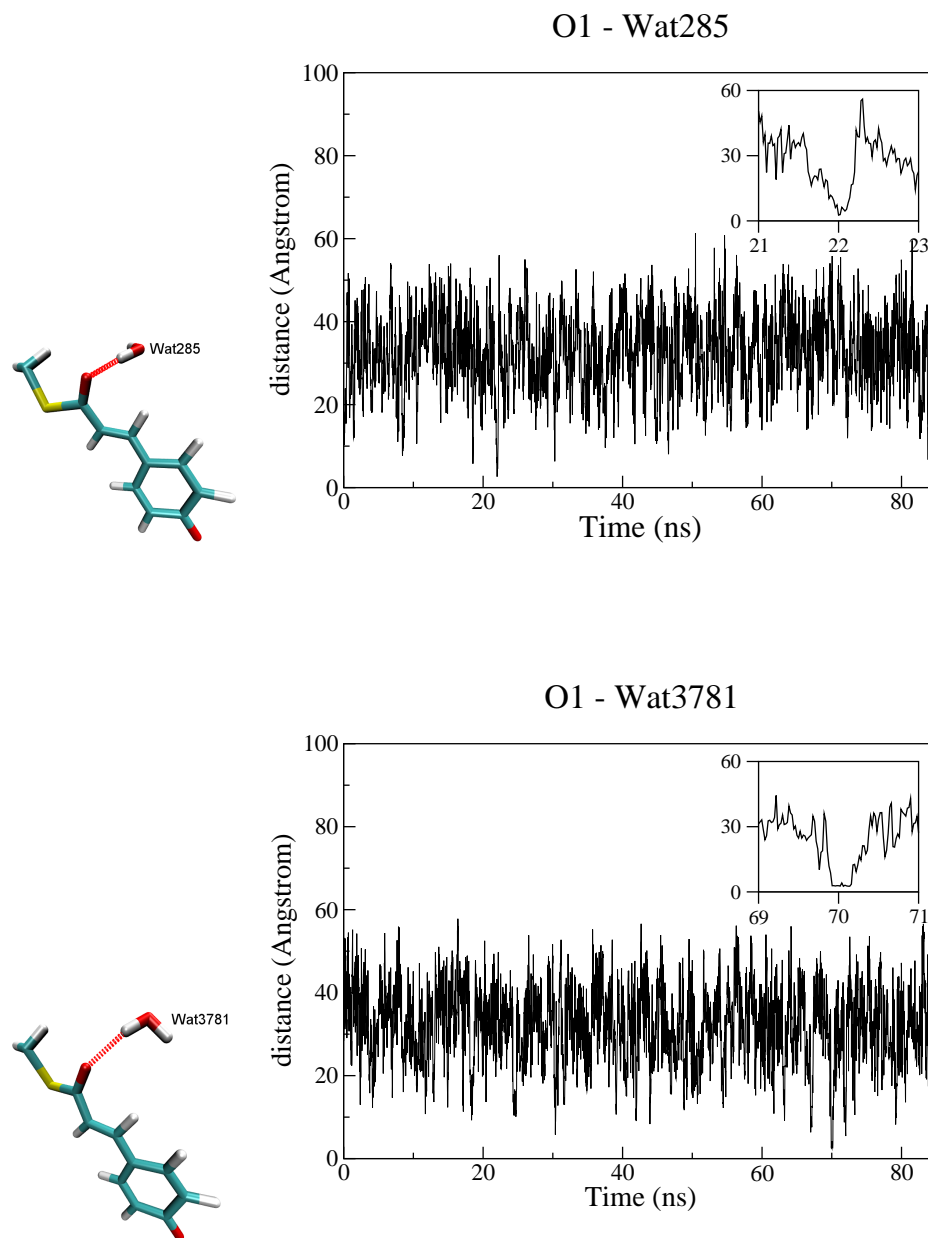


Figure 7.2: Water molecules from the bulk are able to enter the chromophore binding pocket and form hydrogen bonds to the chromophore O1. These hydrogen bonds are stable for mostly less than ten steps whereupon those water molecules move away into the bulk again. Four water molecules (wat274, wat285, wat3781 and wat5625) were randomly chosen to illustrate this occurrence. This figure shows the water molecules wat285 and wat3781 hydrogen bonded to the chromophore O1 and the distances these water molecules have during the MM MD simulation to the O1.

Table 7.4: Isomerisation events for the individual runs which crossed the conical intersection for PYP in water with neutral Arg52. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.

run	surface hop		at surface hop	event
	(step number)	(time in ps)		after surface hop
0	1316	1.316	db	back to trans
6	2877	2.877	sb	back to trans
7	375	0.375	db	back to trans
8	471	0.471	sb	back to trans
13	1063	1.063	db	to cis
14	1324	1.324	sb	
15	465	0.465	sb	back to trans
17	333	0.333	db	back to trans
18	1201	1.201	sb	back to trans
20	893	0.893	sb	back to trans
21	1478	1.478	sb	back to trans
22	846	0.846	sb	
27	189	0.189	db	back to trans
28	760	0.760	sb	back to trans
29	2264	2.264	sb	back to trans
30	1141	1.141	db	back to trans
32	1132	1.132	sb	back to trans
33	1448	1.448	sb	back to trans
35	1084	1.084	sb	back to trans
36	320	0.320	sb	back to trans
38	2580	2.580	sb	
39	1551	1.551	sb	180 degree twist HB to Cys69 broken
40	1486	1.486	sb	back to trans
42	1427	1.427	sb	back to trans
45	691	0.691	sb	back to trans

run	surface hop		event	
	(step number)	(time in ps)	at surface hop	
47	1220	1.220	sb	back to trans
49	3427	3.427	sb	back to trans
52	2830	2.830	db	back to trans
53	1725	1.725	sb	
54	2424	2.424	sb	back to trans
56	495	0.495	sb	
61	201	0.201	sb	back to trans
63	297	0.297	sb	
66	180	0.180	sb	back to trans
69	2904	2.904	sb	
70	2403	2.403	sb	back to trans
71	1151	1.151	sb	back to trans
72	3147	3.147	sb	back to trans
73	1554	1.554	sb	back to trans
75	2528	2.528	db	back to trans
78	2424	2.423	sb	back to trans
79	1597	1.597	sb	back to trans
83	1452	1.452	sb	back to trans
85	2330	2.330	sb	back to trans

Table 7.5: Isomerisation events for the individual runs which crossed the conical Intersection for PYP in water with protonated Arg52. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.

run	surface hop		event	
	(step number)	(time in ps)	at surface hop	after surface hop
0	1098	1.098	sb	back to trans
1	1015	1.015	db	to cis
2	1215	1.215	sb	back to trans
4	945	0.945	sb	back to trans
5	2125	2.125	sb	back to trans
6	124	0.124	db	back to trans
8	1978	1.978	sb	back to trans
12	1716	1.716	db	back to trans
19	1602	1.602	sb	back to trans
20	2767	2.767	sb	back to trans
25	555	0.555	db	back to trans
28	477	0.477	db	to cis
35	2517	2.517	sb	back to trans
36	2711	2.711	sb	back to trans
44	1805	1.805	sb	back to trans
45	1290	1.290	sb	back to trans
48	3024	3.024	db	to cis
51	970	0.970	sb	back to trans
52	861	0.861	db	back to trans
61	2383	2.383	sb	back to trans
66	2150	2.150	db	back to trans
67	2346	2.346	db	to cis
69	2928	2.928	sb	back to trans
75	1187	1.187	db	back to trans
76	2389	2.389	sb	back to trans
86	1554	1.554	sb	back to trans

Influences on the chromophore isomerisation barriers

Table 7.6: Orbitals used for the active space permutations of different sizes of the CASCF active space. The unoccupied orbitals were stable for all permutations while one occupied orbital was exchanged for every calculation. The numbers are taken from Gaussian after the π and π^* orbitals involved in excitation of PYP were chosen as active orbitals.

active space											
12,11	48	47	46	45	44	43	42	41	40	39	38
10,9		47	46	45	44	43	42	41	40	39	
10-9-a		47	46	45	44	43	42	41	40	38	
10-9-b		47	46	45	44	43	42	41	38	39	
10-9-c		47	46	45	44	43	42	38	40	39	
10-9-d		47	46	45	44	43	38	41	40	39	
10-9-e		47	46	45	44	38	42	41	40	39	
8,8		47	46	45	44	43	42	41	40		
8,7			46	45	44	43	42	41	40		
8-7-a			46	45	44	43	42	41	39		
8-7-b			46	45	44	43	42	39	40		
8-7-c			46	45	44	43	39	41	40		
8-7-d			46	45	44	39	42	41	40		
8,6					44	43	42	41			
6,6			46	45	44	43	42	41			
6-6-a			46	45	44	43	42	40			
6-6-b			46	45	44	43	40	41			
6-6-c			46	45	44	40	42	41			

Table 7.7: Barriers calculated with the rigid scan method implemented in Gaussian for all permutations of the CASSCF active space for different sizes of the active space.

active space	db-barrier (kJ/mol)	sb-barrier (kJ/mol)
12-11	24.13	2.625
10-10	17.5	10.8
10-9	16.25	0.604
10-9-a	18.98	2.94
10-9-b	18.98	2.94
10-9-c	17.75	—
10-9-d	18.98	2.94
10-9-e	22.45	—
8-8	10.7	5.4
8-7	11.74	0.866
8-7-a	11.74	0.866
8-7-b	11.76	0.893
8-7-c	11.74	0.866
8-7-d	11.74	0.866
8-6	7.93	—
6-6	35.5	—
6-6-a	3.6	5.43
6-6-b	29.12	—
6-6-c	15.46	—

Table 7.8: Isomerisation barriers published by different authors

publication	double bond barrier (kJ/mol)	single bond barrier (kJ/mol)
Ko <i>et al.</i> (2008) Chem. Phys. Letters (CASSCF)		
positive pointcharge	no barrier	2.89
neutral pointcharge	6.75	1.93
Coto <i>et al.</i> (2009) JCTC (CASSCF)	14.47	
de Visser <i>et al.</i> (pers. correspondence) (CASSCF)	11.89	
Gromov <i>et al.</i> (pers. correspondence) (CC2)	25.09	

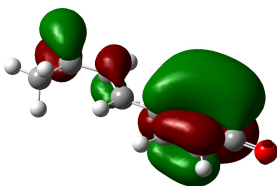
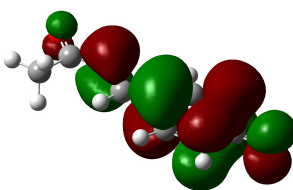
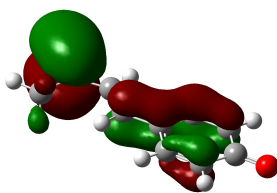
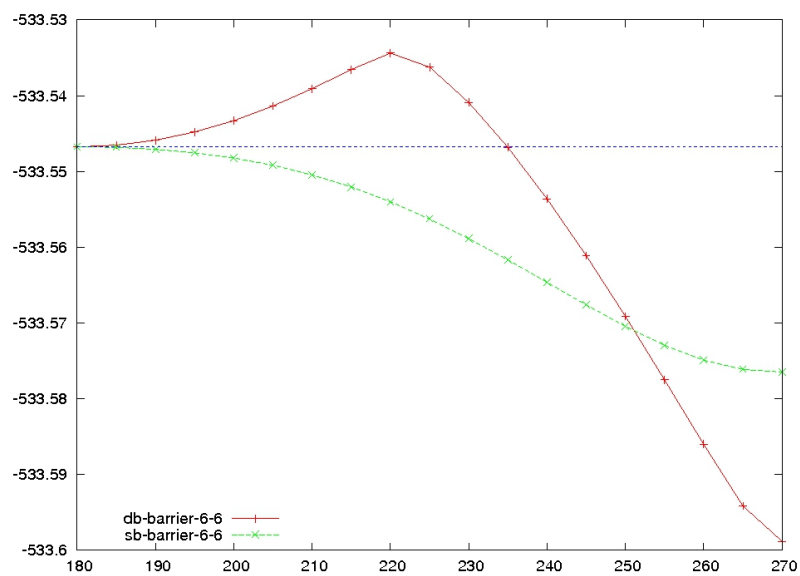


Figure 7.3: Isomerisation barriers and bound orbitals of permutation 6-6

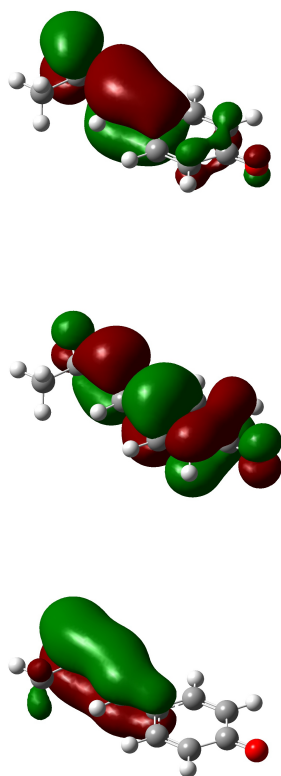
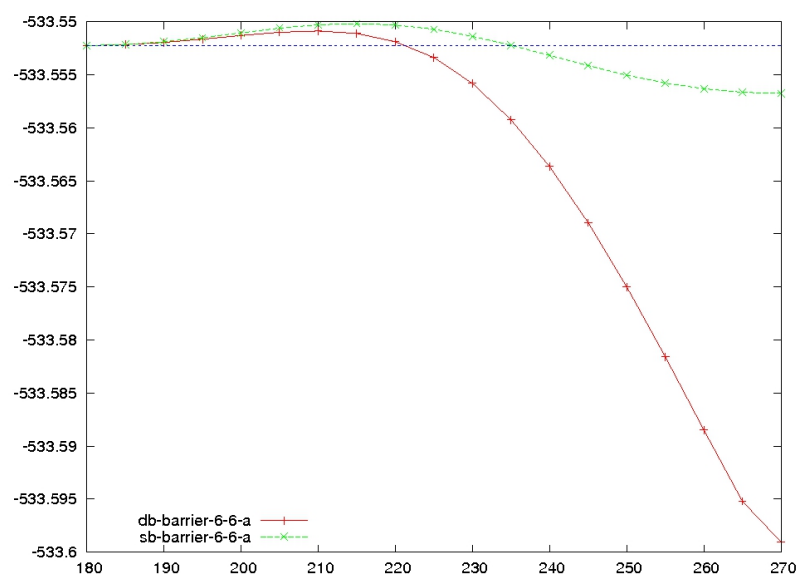


Figure 7.4: Isomerisation barriers and bound orbitals of permutation 6-6-a

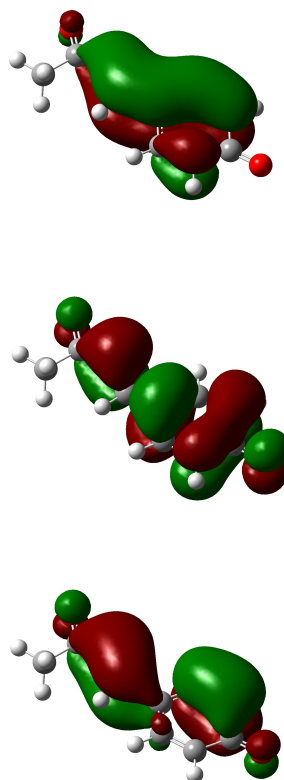
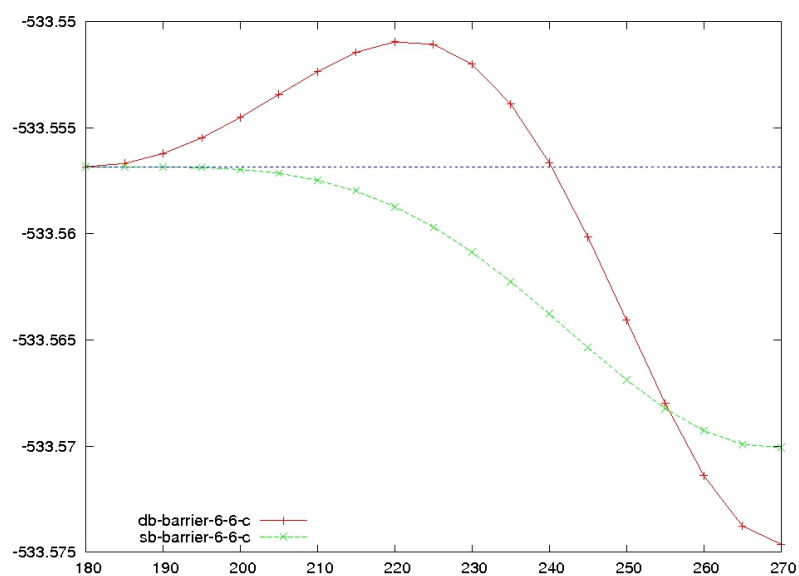


Figure 7.5: Isomerisation barriers and bound orbitals of permutation 6-6-c

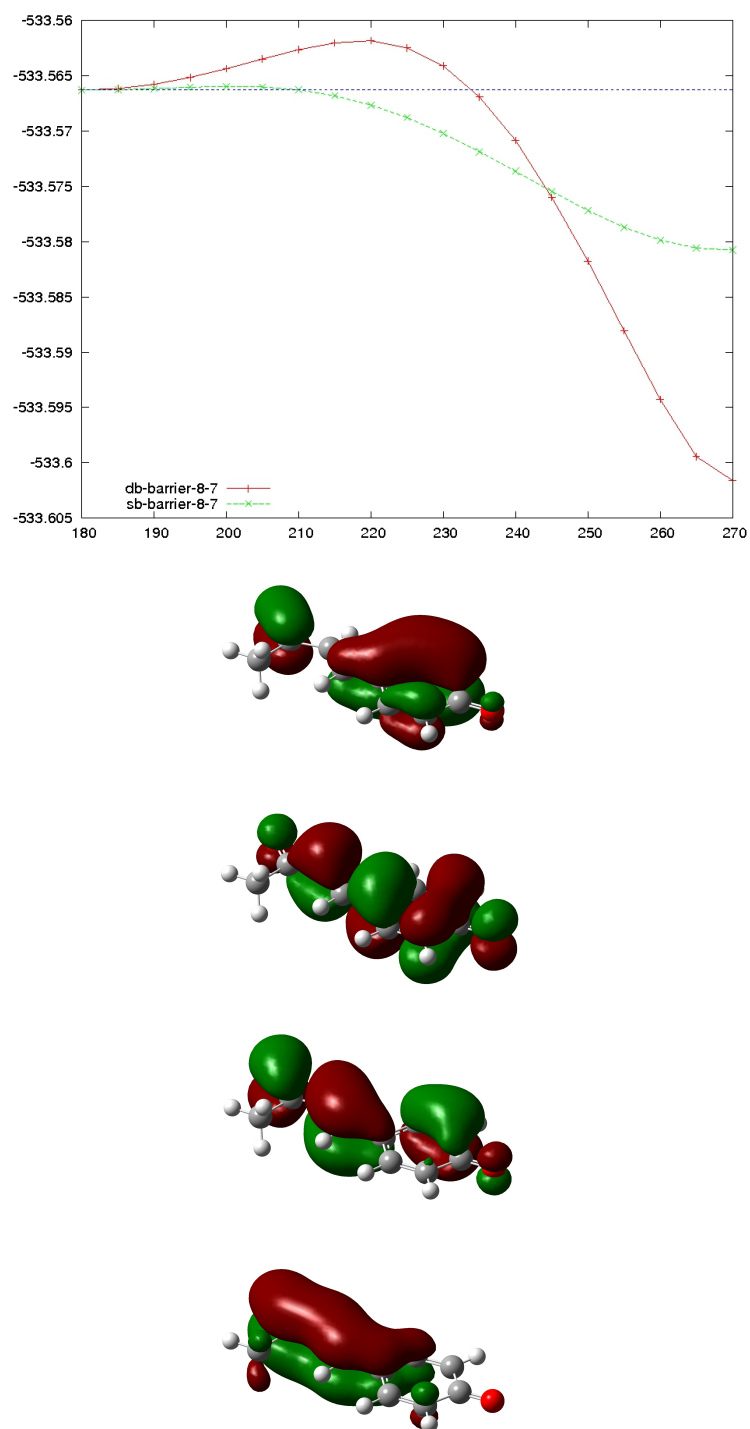


Figure 7.6: Isomerisation barriers and bound orbitals of permutation 8-7-a

Isomerisation events of pCK⁻ in water

Table 7.9: Summary of moment of surface hop and corresponding deactivation events for pCK⁻ in water. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.

run	surface hop		event	
	(step number)	(time in ps)		
2	1466	0.733	sb	back to gs
3	1692	0.846	sb	180 degree turn
4	219	0.1095	db	to cis
5	883	0.4415	db	back to gs
6	615	0.3075	sb	back to gs
7	2490	1.245	sb	back to gs
8	2473	1.2365	sb	back to gs
9	750	0.375	sb	180 degree turn
10	2326	1.163	db	to cis
11	1320	0.66	sb	back to gs
12	1599	0.7995	sb	180 degree turn
14	1452	0.726	sb	180 degree turn
15	1799	0.8995	sb	180 degree turn
16	1797	0.8985	sb	back to gs
17	608	0.304	sb	back to gs
18	165	0.0825	db	back to gs
19	1690	0.845	sb	back to gs
20	3929	1.9645	sb	back to gs
22	2133	1.0665	sb	back to gs
23	1316	0.658	sb	back to gs
24	1626	0.813	sb	back to gs
26	684	0.342	sb	back to gs
27	2135	1.0675	sb	back to gs
28	443	0.2215	sb	back to gs
29	264	0.132	db	to cis
30	358	0.179	db	to cis
31	301	0.1505	db	back to gs
32	3684	1.842	db	back to gs
33	619	0.3095	sb	180 degree turn

run	surface hop		event	
	(step number)	(time in ps)	at surface hop	after surface hop
34	1141	0.5705	sb	
35	188	0.094	db	back to gs
36	238	0.119	db	back to gs
37	2660	1.33	sb	180 degree turn
38	837	0.4185	sb	back to gs
39	250	0.125	db	back to gs
40	1676	0.838	sb	back to gs
41	917	0.4585	sb	back to gs
42	2229	1.1145	sb	back to gs
43	449	0.2245	sb	180 degree turn
44	1215	0.6075	sb	back to gs
46	785	0.3925	sb	back to gs
48	2351	1.1755	sb	back to gs
49	2702	1.351	sb	
50	1555	0.7775	sb	back to gs
51	485	0.2425	db	back to gs
54	1689	0.8445	sb	back to gs
55	1561	0.7805	sb	
56	1346	0.673	sb	back to gs
58	411	0.2055	db	back to gs
60	669	0.3345	db	to cis
64	1467	0.7335	sb	back to gs
66	722	0.361	db	to cis
68	239	0.1195	db	back to gs

List of Figures

1.1	Structure of PYP with the position of the chromophore bound to Cys69 via a thio-ester linkage (depicted as explicit atoms). The β -sheet is shown in yellow whereas the connecting α -helices are coloured purple.	2
1.2	The PYP chromophore surrounded by the amino acids that form the binding pocket. The chromophore forms hydrogen bonds to Tyr42, Glu46, Thr50, and Cys69. Arg52 and Thr50 and Tyr98 are connected by a hydrogen bond among themselves.	3
1.3	The PYP photocycle with the ground state (gs, pG), the excited state (es) after photon absorption by the chromophore, the pR state after deactivation of the chromophore from the excited state via db isomerisation, and the signalling state (pB), where protonation and structural changes are observed.	4
1.4	The planar chromophore in the ground state and the twisted structures resulting from the two possible deactivation pathways: double bond isomerisation (1) and single bond rotation (2)	7
1.5	Asymmetrical hydrogen bond potential for a typical hydrogen bond. The two electronegative atoms have different chemical properties and the donor-acceptor distance is 2.7–3 Å. The hydrogen is covalently bound to the hydrogen donor.. . . .	9
1.6	A low barrier hydrogen bond (LBHB) potential has two equal or almost equal minima along the hydrogen bond. The barrier separating them is low enough for the hydrogen to occupy both forming covalent bonds to both donor and acceptor. This symmetrical double well potential arises from similar electronegativity of hydrogen donor and acceptor and shorter donor-acceptor distance.	10

1.7	In a single well hydrogen bond potential no energy barrier is present along the bond between hydrogen donor and acceptor. The hydrogen is centred between both electronegative atoms. Donor and acceptor have almost the same electronegativity and the distance between them is less than 2.5 Å.	10
2.1	Scheme of QM/MM	20
2.2	Differences between RHF, full CI, and CASSCF. For RHF, each orbital can be either doubly occupied or unoccupied. Full CI has no restrictions regarding the orbital occupancy and all possible excitations are allowed. The combination of both approaches is CASSCF, where the active orbitals are treated as in the case of full CI. Orbitals with lower energy than the active orbitals are always doubly occupied, and orbitals with higher energy are always unoccupied as it would be the case for RHF.	32
2.3	Chromophore model pCK ⁻ which is a ketone derivative of the deprotonated <i>trans</i> - <i>p</i> -coumaric acid.	34
2.4	The three highest occupied molecular orbitals of the chromophore model pCK ⁻ for the CAS 6,6 active space permutations 6-6 (1), 6-6-a (2) and 6-6-c (3). The phenyl ring of the chromophore is always on the lower right side of each image and the ethylene chain on the upper left side.	36
2.5	The potential energy surfaces of the excited state S1 (red) and the ground state S0 (blue) along the <i>trans</i> -to- <i>cis</i> isomerisation coordinate <i>b</i> and a skeletal deformation of the bonds. The seam between the two surfaces consists of conical intersections. A typical trajectory is depicted in as yellow line. It starts on the ground state before excitation. On the excited state, it moves from the Franck-Condon-region over a barrier to the db-twisted minimum. From this geometry a hop to the ground state via a conical intersection takes place. On ground state evolution to <i>trans</i> - and <i>cis</i> -conformation are possible.(Adapted with permission from Groenhof <i>et al.</i> ²⁵ . Copyright 2004 American Chemical Society.)	39

3.1	Comparison of the energy potential with 1.0 nm (black) and 2.5 nm cut-off (red) for structure A with QM region 3 (see table 3.1) and neutral Arg52. A larger cut-off which considers the electrostatic effects of the whole protein unto the QM region changes the shape of the energy potential significantly. Not taking into account parts of the protein results in too large barrier heights.	45
3.2	Comparison of the energy potential of calculations using ONIOM and the normal QMMM scheme for minimised structure A with QM region 3 and protonated/unprotonated Arg52.	46
3.3	The energy potentials of structures A (black) and B (red) for QM region 3 with neutral Arg52. The differing orientations of seven amino acids side chains in structure B compared to structure A cause a preference of the delocalised proton.	48
3.4	Comparison of the energy potential of structures A, B and single differing amino acids for QM region 3 with neutral Arg52. Structure A and B differ in the positions of seven amino acid side chains. When the different coordinates from structure B for one of these amino acid side chains is used in structure A only for Lys123 the different potential energy could be observed.	49
3.5	Energy potentials along the hydrogen bond of all six QM regions specified in Table 3.1 for protonated (up) and neutral (down) Arg52. Structure A and a cut off radius of 2.5 nm were used to calculate the potentials.	52
3.6	The energy potentials (red), wave functions (black), probability densities (green) and expectation values (blue) along the hydrogen bond between the chromophore and Glu46 for neutral (up) and protonated Arg52 (down).	54
3.7	Comparison of the energy potential of minimised (red) and unminimised (black) structure A for QM region 3 with neutral Arg52. Minimising the system results in the same qualitative picture with a larger barrier height and less pronounced minimum at the chromophore.	56
4.1	Illustration of the amino acids forming the chromophore binding pocket that are relevant in this work with identification of the atoms that are specifically mentioned in the descriptions of this chapter.	63

-
- 4.2 Comparison of the minimised structure with protonated Arg52 (blue) with the X-ray structure 2ZOI (red) on the left side. The right picture depicts the minimised structure (blue) and a snapshot from the MM MD simulation (orange) both with protonated Arg52. The chromophore binding pocket is conserved by the minimisation and the MM MD simulation. . . . 65
- 4.3 Depiction of the RMSD for the backbone (black) and the whole protein except the chromophore (red) throughout the MM MD simulation with protonated Arg52. The minimised structure was used as the reference structure. 66
- 4.4 During the MM MD simulation with protonated Arg52 the length of the hydrogen bonds between the chromophore O4a and Tyr42 (black) and Glu46 (red) evolved around 2.8 and 2.7 Å, respectively. 67
- 4.5 Positions of the chromophore, Tyr42, Glu46, Thr50, Arg52, Cys69 and Tyr98 in the initial structure of the MM MD simulation. The positions correspond to the positions observed in the crystal structure 2ZOH. 68
- 4.6 Comparison of Arg52 movement for MM MD with protonated (black) and neutral Arg52 (red) by measuring the distance between Arg52 NH and Tyr98 backbone O. For protonated Arg52 the distance between the Arg52 side chain and the Tyr98 backbone is constant at 2.75 Å in average. The hydrogen bond between both residues is intact throughout the simulation. This hydrogen bond does not exist for the simulation with deprotonated Arg52. Its side chain rotates away from Tyr98 into the bulk water and the hydrogen bond is broken. 69
- 4.7 The number of hydrogen bonds between the chromophore's O1 and water molecules from the bulk for every 20th step of the simulation of PYP in water with neutral Arg52. Water molecules are able to enter the chromophore binding pocket because of the Arg52 side chain movement. Up to two water molecules are found bound to O1 at a time. 70

-
- 4.8 Water molecules from the bulk are able to enter the chromophore binding pocket and form hydrogen bonds to the chromophore O1. These hydrogen bonds are stable for mostly less than ten steps whereupon the water molecules move away into the bulk again. Four water molecules (Wat274, wat285, wat3781 and wat5625) were randomly chosen to illustrate this occurrences. This figure shows the water molecules wat274 and wat5625 hydrogen bonded to the chromophore O1 and the distances these water molecules have during the MM MD simulation to the O1. The data for the remaining two water molecules is shown in the appendix (figure 7.2). 71
- 4.9 Structure of the chromophore binding pocket and surrounding amino acids after the hydrogen bond between Arg52 and Tyr98 has been broken and side chains of both amino acids moved into their new position. The water molecule bridging Thr50 and the chromophore was able to enter the binding pocket because of the before described movement. 72
- 4.10 Chromophore binding pocket during the movement of the chromophore. The hydrogen bonds between the chromophore and Tyr42 and Glu46 are preserved by a bridging water molecule which was able to enter the chromophore binding pocket because of the broken hydrogen bond between Arg52 and Tyr98. 73
- 4.11 Length of hydrogen bonds between the chromophore and Tyr42 (black) and Glu46 (red) during MM MD simulations with deprotonated Arg52. The hydrogen bonds are broken between step 430 and 730 because the chromophore moves out of the binding pocket. Afterwards the initial conformation is re-established for the remaining simulation time. 74
- 4.12 Comparison of the RMSF values of the individual PYP residues for protonated (red) and neutral Arg52 (black). Residues 52 and 98 and the neighbouring amino acids show the largest difference between both protonation states due to the large displacements observed in the MM MD simulations with neutral Arg52. Fluctuations of Ser114 are observed because this residue, located at the protein surface, does not form hydrogen bonds to other parts of the protein and therefore its side chain is able to move freely. 75

4.13	Hydrogen bond length between (A) the chromophore's O4a and Glu46 and (B) the chromophore's O4a and Tyr42 for QMMM MD simulations with protonated (black) and neutral Arg52 (red). For both protonated and neutral Arg52 the hydrogen bond length is in average the same.	77
4.14	Distance between NH of Arg52 and the backbone O of Tyr98 for QMMM MD simulations with protonated (+) and neutral (x) Arg52. For neutral Arg52 the distance between the Arg52 side chain and the Tyr98 backbone is larger than for protonated Arg52. The hydrogen bond existing in case of protonated Arg52 is not formed in case of neutral Arg52.	78
4.15	Evolution of the hydrogen bond length between O4a and Tyr42 (left) and Glu46 (right) during QMMM MD simulations with (red) and without (black) diffuse functions for PYP with protonated Arg52. The usage of diffuse functions did not provoke shorter or longer hydrogen bonds compared to not using diffuse functions.	79
4.16	Evolution of the hydrogen bond length between O4a and Tyr42 (left) and Glu46 (right) during QMMM MD simulations with (red) and without (black) diffuse functions for PYP with neutral Arg52. The usage of diffuse functions did not provoke shorter or longer hydrogen bonds compared to not using diffuse functions.	80
4.17	Depiction of strained geometry of the chromophore's ethylene chain after deactivation from excited state to ground state in <i>cis</i> -conformation. (1) shows the even distribution over the three affected bonds whereas the twist is not dispersed over the double bond in (2).	87
4.18	Depiction of the hydrogen bond between the chromophore O1 and the Cys69 N after deactivation from the excited state to <i>cis</i> -conformation on the ground state and subsequent QMMM MD simulation. The left picture shows the intact hydrogen bond and a strained geometry of the chromophore while the picture on the right displays the broken hydrogen bond with the unstrained chromophore.	88
4.19	Charge localisation on the chromophore for ground state (1) and excited state (2). In ground state the negative charge is localised on O4a. Through light excitation a charge transfer towards the ethylene chain and O1 occurs.	94
4.20	Chromophore with the amino acids Tyr42, Glu46 and Cys69	96

4.21	Single (red) and double bond (black) isomerisation barrier heights for hydrogen bond length of 2.90Å each. The db barrier is significantly higher than the sb barrier.	97
4.22	Single (re) and double bond (black) isomerisation barrier heights for hydrogen bond length taken from X-ray structure 2ZOH. Both barriers are of about the same heights.	98
4.23	In excited QMMM simulations twisted intermediates were observed after deactivation via db isomerisation and subsequent rotation to the <i>cis</i> -conformation. Similar conformations were reported recently by Schotte <i>et al.</i> and Jung <i>et al.</i> through time-resolved Laue-crystallography. The theoretical results presented in this thesis match the conformation shown by the experimental studies.	102
7.1	Depiction of the planar <i>trans</i> -configured chromophore with atom names.	125
7.2	Water molecules from the bulk are able to enter the chromophore binding pocket and form hydrogen bonds to the chromophore O1. These hydrogen bonds are stable for mostly less than ten steps whereupon thoe water molecules move away into the bulk again. Four water molecules (wat274, wat285, wat3781 and wat5625) were randomly chosen to illustrate this occurences. This figure shows the water molecules wat285 and wat3781 hydrogen bonded to the chromophore O1 and the distances these water molecules have during the MM MD simulation to the O1.	129
7.3	Isomerisation barriers and bound orbitals of permutation 6-6 .	135
7.4	Isomerisation barriers and bound orbitals of permutation 6-6-a	136
7.5	Isomerisation barriers and bound orbitals of permutation 6-6-c	137
7.6	Isomerisation barriers and bound orbitals of permutation 8-7-a	138

List of Tables

2.1	Barriers for db isomerisation and sb rotation calculated with the rigid scan method implemented in Gaussian.	34
2.2	Db isomerisation and sb rotation barriers for permuted CAS 6,6 active spaces.	34
3.1	Numbering and selected amino acids of the different QM regions for which the potential energy curve was calculated to compare the effect of choosing different compositions of the QM region	50
4.1	Observed torsions for excited states QMMM MD simulations. The values for simulations which showed deactivation during the simulation time are written in brackets. Simulations with neutral Arg52 show more deactivations than simulations with protonated Arg52. In the latter case proportionally more deactivations took place via db isomerisation. Neutral Arg52 favoured deactivation via sb rotation.	86
4.2	Average time until deactivation via both pathways from the excited state for different Arg52 protonation states. The deactivation via db isomerisation is faster than via sb rotation. Deactivations happen earlier for the case of neutral Arg52. . .	86
4.3	Events on ground state after deactivations from the excited state via db isomerisation. Of five simulations which deactivated to the <i>cis</i> -conformation on the ground state four maintain this conformation. For two of these the hydrogen bond to Cys69 is broken during the ground state QMMM MD simulations.	87
4.4	Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for neutral and protonated Arg52	89

4.5	Number of hydrogen bonds at the moment of excitation for all individual simulations with neutral Arg52 that showed deactivation from the excited state to the ground state.	91
4.6	Number of hydrogen bonds at the moment of excitation for all individual simulations with protonated Arg52 that showed deactivation from the excited state to the ground state.	92
4.7	Comparison of the estimated barrier heights for single and double bond isomerisation for different hydrogen bond length. The shorter the hydrogen bonds to Tyr42 and Glu46 are, the higher the sb barrier and the lower the db barrier.	95
5.1	Summary of the simulation results for the chromophore model pCK ⁻ in water. No twists could be observed for 12 simulations, which corresponds to 17.65 %.	110
5.2	Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for pCK ⁻ in water.	110
5.3	Summary of the simulation results for the chromophore model pCK ⁻ in decanol. No twists could be observed for 33 simulations, which corresponds to 34.375 %.	112
5.4	Average number of hydrogen bonds at the moment of excitation for the two possible deactivation pathways for pCK ⁻ in decanol.	112
7.1	Amber atom type and charge for the PYP chromophore parametrised by MP2 geometry optimisation and subsequent determination of the charges by restricted electrostatic potential fitting (RESP).	126
7.2	Bond length between the atoms of the PYP chromophore for use in the AMBER03 force field calculated by MP2 optimisation	127
7.3	Angles present in the PYP chromophore for use in the AMBER03 force field calculated by MP2 optimisation	128
7.4	Isomerisation events for the individual runs which crossed the conical intersection for PYP in water with neutral Arg52. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.	130

7.5	Isomerisation events for the individual runs which crossed the conical Intersection for PYP in water with protonated Arg52. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.	132
7.6	Orbitals used for the active space permutations of different sizes of the CASCF active space. The unoccupied orbitals were stable for all permutations while one occupied orbital was exchanged for every calculation. The numbers are taken from Gaussian after the π and π^* orbitals involved in excitation of PYP were chosen as active orbitals.	133
7.7	Barriers calculated with the rigid scan method implemented in Gaussian for all permutations of the CASSCF active space for different sizes of the active space.	134
7.8	Isomerisation barriers published by different authors	134
7.9	Summary of moment of surface hop and corresponding deactivation events for pCK ⁻ in water. The table depicts for all runs that deactivated from excited state to ground state the step and corresponding time at which the deactivation took place, the occurred deactivation event and the resulting structure on gs.	139

Bibliography

- [1] T. E. Meyer, *BBA - Bioenergetics* **806**(1), 175–183 (1985).
- [2] W. W. Sprenger, W. D. Hoff, J. P. Armitage, and K. J. Hellingwerf, *Journal of Bacteriology* **175**(10), 3096–3104 (1993).
- [3] M. Baca, G. E. O. Borgstahl, M. Boissinot, P. M. Burke, D. R. Williams, K. A. Slater, and E. D. Getzoff, *Biochemistry* **33**(48), 14369–14377 (1994).
- [4] C. P. Ponting and L. Aravind, *Current biology : CB* **7**(11), R674–677 (1997).
- [5] J. . Pellequer, K. A. Wager-Smith, S. A. Kay, and E. D. Getzoff, *Proceedings of the National Academy of Sciences of the United States of America* **95**(11), 5884–5890 (1998).
- [6] D. M. Lagarias, Wu Shu-Hsing, and J. C. Lagarias, *Plant Molecular Biology* **29**(6), 1127–1142 (1995).
- [7] M. H. Hefti, K. . François, S. C. De Vries, R. Dixon, and J. Vervoort, *European Journal of Biochemistry* **271**(6), 1198–1208 (2004).
- [8] B. L. Taylor and I. B. Zhulin, *Microbiology and Molecular Biology Reviews* **63**(2), 479–506 (1999).
- [9] M. A. Cusanovich and T. E. Meyer, *Biochemistry* **42**(17), 4759–4770 (2003).
- [10] Y. Imamoto and M. Kataoka, *Photochemistry and photobiology* **83**(1), 40–49 (2007).
- [11] J. T. Henry and S. Crosson, **65**, 261–286 (2011).
- [12] J. J. Van Beeumen, B. V. Devreese, S. M. Van Bun, W. D. Hoff, K. J. Hellingwerf, T. E. Meyer, D. E. McRee, and M. A. Cusanovich, *Protein Science* **2**(7), 1114–1125 (1993).

- [13] W. D. Hoff, I. H. M. Van Stokkum, H. J. Van Ramesdonk, M. E. Van Brederode, A. M. Brouwer, J. C. Fitch, T. E. Meyer, R. Van Grondelle, and K. J. Hellingwerf, *Biophysical journal* **67**(4), 1691–1705 (1994).
- [14] M. Kim, R. A. Mathies, W. D. Hoff, and K. J. Hellingwerf, *Biochemistry* **34**(39), 12669–12672 (1995).
- [15] T. E. Meyer, E. Yakali, M. A. Cusanovich, and G. Tollin, *Biochemistry* **26**(2), 418–423 (1987).
- [16] A. Xie, L. Kelemen, J. Hendriks, B. J. White, K. J. Hellingwerf, and W. D. Hoff, *Biochemistry* **40**(6), 1510–1517 (2001).
- [17] M. Yoda, Y. Inoue, and M. Sakurai, *Journal of Physical Chemistry B* **107**(51), 14569–14575 (2003).
- [18] U. K. Genick, G. E. O. Borgstahl, K. Ng, Z. Ren, C. Pradervand, P. M. Burke, V. Šrajer, T. . Teng, W. Schildkamp, D. E. McRee, K. Moffat, and E. D. Getzoff, *Science* **275**(5305), 1471–1475 (1997).
- [19] R. Brudler, T. E. Meyer, U. K. Genick, S. Devanathan, T. T. Woo, D. P. Millar, K. Gerwert, M. A. Cusanovich, G. Tollin, and E. D. Getzoff, *Biochemistry* **39**(44), 13478–13486 (2000).
- [20] N. Shimizu, H. Kamikubo, K. Mihara, Y. Imamoto, and M. Kataoka, *Journal of Biochemistry* **132**(2), 257–263 (2002).
- [21] N. Shimizu, H. Kamikubo, Y. Yamazaki, Y. Imamoto, and M. Kataoka, *Biochemistry* **45**(11), 3542–3547 (2006).
- [22] U. K. Genick, S. M. Soltis, P. Kuhn, I. L. Canestrelli, and E. D. Getzoff, *Nature* **392**(6672), 206–209 (1998).
- [23] B. Perman, V. Šrajer, Z. Ren, T. . Teng, C. Pradervand, T. Ursby, D. Bourgeois, F. Schotte, M. Wulff, R. Kort, K. Hellingwerf, and K. Moffat, *Science* **279**(5358), 1946–1950 (1998).
- [24] Z. Ren, B. Perman, V. Šrajer, T. . Teng, C. Pradervand, D. Bourgeois, F. Schotte, T. Ursby, R. Kort, M. Wulff, and K. Moffat, *Biochemistry* **40**(46), 13788–13801 (2001).
- [25] G. Groenhof, M. Bouxin-Cademartory, B. Hess, S. P. De Visser, H. J. C. Berendsen, M. Olivucci, A. E. Mark, and M. A. Robb, *Journal of the American Chemical Society* **126**(13), 4228–4233 (2004).

- [26] E. V. Gromov, I. Burghardt, J. T. Hynes, H. Köppel, and L. S. Cederbaum, *Journal of Photochemistry and Photobiology A: Chemistry* **190**(2-3), 241–257 (2007).
- [27] M. Boggio-Pasqua, M. A. Robb, and G. Groenhof, *Journal of the American Chemical Society* **131**(38), 13580–13581 (2009).
- [28] A. M. Virshup, C. Punwong, T. V. Pogorelov, B. A. Lindquist, C. Ko, and T. J. Martínez, *Journal of Physical Chemistry B* **113**(11), 3280–3291 (2009).
- [29] M. Boggio-Pasqua and G. Groenhof, *Journal of Physical Chemistry B* **115**(21), 7021–7028 (2011).
- [30] P. Changenet-Barret, A. Espagne, S. Charier, J. . Baudin, L. Jullien, P. Plaza, K. J. Hellingwerf, and M. M. Martin, *Photochemical and Photobiological Sciences* **3**(8), 823–829 (2004).
- [31] P. Changenet-Barret, A. Espagne, P. Plaza, K. J. Hellingwerf, and M. M. Martin, *New Journal of Chemistry* **29**(4), 527–534 (2005).
- [32] A. Espagne, P. Changenet-Barret, P. Plaza, and M. M. Martin, *Journal of Physical Chemistry A* **110**(10), 3393–3404 (2006).
- [33] E. V. Gromov, I. Burghardt, H. Köppel, and L. S. Cederbaum, *Journal of Physical Chemistry A* **115**(33), 9237–9248 (2011).
- [34] P. Changenet-Barret, P. Plaza, M. M. Martin, H. Chosrowjan, S. Taniguchi, N. Mataga, Y. Imamoto, and M. Kataoka, *Chemical Physics Letters* **434**(4-6), 320–325 (2007).
- [35] P. Changenet-Barret, P. Plaza, M. M. Martin, H. Chosrowjan, S. Taniguchi, N. Mataga, Y. Imamoto, and M. Kataoka, *Journal of Physical Chemistry C* **113**(27), 11605–11613 (2009).
- [36] E. V. Gromov, I. Burghardt, H. Köppel, and L. S. Cederbaum, *Journal of the American Chemical Society* **129**(21), 6798–6806 (2007).
- [37] G. Groenhof, M. F. Lensink, H. J. C. Berendsen, J. G. Snijders, and A. E. Mark, *Proteins: Structure, Function and Genetics* **48**(2), 202–211 (2002).
- [38] L. L. Premvardhan, M. A. Van der Horst, K. J. Hellingwerf, and R. Van Grondelle, *Biophysical journal* **84**(5), 3226–3239 (2003).

-
- [39] A. Xie, W. D. Hoff, A. R. Kroon, and K. J. Hellingwerf, *Biochemistry* **35**(47), 14671–14678 (1996).
- [40] M. L. Groot, L. J. G. W. Van Wilderen, D. S. Larsen, M. A. Van der Horst, I. H. M. Van Stokkum, K. J. Hellingwerf, and R. Van Grondelle, *Biochemistry* **42**(34), 10054–10059 (2003).
- [41] L. L. Premvardhan, F. Buda, M. A. Van Der Horst, D. C. Lührs, K. J. Hellingwerf, and R. Van Grondelle, *Journal of Physical Chemistry B* **108**(16), 5138–5148 (2004).
- [42] U. K. Genick, S. Devanathan, T. E. Meyer, I. L. Canestrelli, E. Williams, M. A. Cusanovich, G. Tollin, and E. D. Getzoff, *Biochemistry* **36**(1), 8–14 (1997).
- [43] G. C. Pimentel and A. L. McClellan, *Annual Review of Physical Chemistry* **22**, 347–385 (1971).
- [44] P. A. Kollman and L. C. Allen, *Chemical reviews* **72**(3), 283–303 (1972).
- [45] W. W. Cleland, P. A. Frey, and J. A. Gerlt, *Journal of Biological Chemistry* **273**(40), 25529–25532 (1998).
- [46] Y. Noda, H. Kasatani, Y. Watanabe, and H. Terauchi, *Journal of the Physical Society of Japan* **61**(3), 905–915 (1992).
- [47] F. Hibbert and J. Emsley, *Advances in Physical Organic Chemistry* **26**, 255–379 (1990).
- [48] T. H. Steiner and W. Saenger, *Acta Crystallographica Section B* **50**(3), 348–357 (1994).
- [49] S. Yamaguchi, H. Kamikubo, K. Kurihara, R. Kuroki, N. Niimura, N. Shimizu, Y. Yamazaki, and M. Kataoka, *Proceedings of the National Academy of Sciences of the United States of America* **106**(2), 440–444 (2009).
- [50] P. Gilli, V. Bertolasi, V. Ferretti, and G. Gilli, *Journal of the American Chemical Society* **116**(3), 909–915 (1994).
- [51] W. R. Davidson, J. Sunner, and P. Kebarle, *Journal of the American Chemical Society* **101**(7), 1675–1680 (1979).
- [52] M. Meot-Ner, *Journal of the American Chemical Society* **106**(5), 1257–1264 (1984).

- [53] J. W. Larson and T. B. McMahon, *Journal of the American Chemical Society* **105**(10), 2944–2950 (1983).
- [54] C. L. Perrin and J. B. Nielson, *"Strong" hydrogen bonds in chemistry and biology* volume 48 of *Annual Review of Physical Chemistry*, (1997).
- [55] C. L. Perrin, *Accounts of Chemical Research* **43**(12), 1550–1557 (2010).
- [56] J. P. Guthrie, *Chemistry and Biology* **3**(3), 163–170 (1996).
- [57] A. Warshel, A. Papazyan, and P. A. Kollman, *Science* **269**(5220), 102–106 (1995).
- [58] A. Warshel and A. Papazyan, *Proceedings of the National Academy of Sciences of the United States of America* **93**(24), 13665–13670 (1996).
- [59] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl, *Journal of Chemical Theory and Computation* **4**(3), 435–447 (2008).
- [60] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, "Gaussian 03, Revision D.01".
- [61] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro,

- M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, "Gaussian 09 Revision A.02".
- [62] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, "Molpro, version 2012.1, a package of ab initio programs".
- [63] William Humphrey, Andrew Dalke, and Klaus Schulten, *Journal of Molecular Graphics* **14**, 33–38 (1996).
- [64] John Stone, "an efficient library for parallel ray tracing and animation", *Master's thesis Computer Science Department, University of Missouri-Rolla (1998)*.
- [65] F. Jensen, *Introduction to Computational Chemistry*, Wiley Chichester (2007).
- [66] C. J. Cramer, *Essentials of Computational Chemistry*, Wiley Chichester (2004).
- [67] P.-O. Widmark (Ed.), *European Summerschool in Quantum Chemistry 2011*, Lund University Lund (2011).
- [68] B. Frenkel, D.; Smit, *Understanding Molecular Simulation*, Academic Press San Diego 2 Auflage (2002).
- [69] T. Schlick, *Molecular Modeling and Simulation*, Springer New York (2006).
- [70] L. Verlet, *Phys. Rev.* **159**(1), 98–103 (1967).

-
- [71] *W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson*, J. Chem. Phys. **76**(1), 637–649 (1981).
- [72] *M. Tuckerman, B. J. Berne, and G. J. Martyna*, J. Chem. Phys. **97**(3), 1990–2001 (1992).
- [73] *H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak*, J. Chem. Phys. **81**(8), 3684–3690 (1984).
- [74] *G. Bussi, D. Donadio, and M. Parrinello*, Journal of Chemical Physics **126**(1) (2007).
- [75] *J. E. Lennard-Jones*, Proc. Phys. Soc. **43**(5), 461–482 (1931).
- [76] *D. Riccardi, P. Schaefer, Y. Yang, H. Yu, N. Ghosh, X. Prat-Resina, P. König, G. Li, D. Xu, H. Guo, M. Elstner, and Q. Cui*, J. Phys. Chem. B **110**(13), 6458–6469 (2006).
- [77] *H. M. Senn and W. Thiel*, Curr. Opin. Chem. Biol. **11**(2), 182–187 (2007).
- [78] *P. Hohenberg and W. Kohn*, Physical Review **136**(3B), B864–B871 (1964).
- [79] *W. Kohn and L. J. Sham*, Physical Review **140**(4A), A1133–A1138 (1965).
- [80] *K. K. Baeck and T. J. Martinez*, Chemical Physics Letters **375**(3-4), 299–308 (2003).
- [81] *M. Sulpizi, U. F. Röhrig, J. Hutter, and U. Rothlisberger*, International Journal of Quantum Chemistry **101**(6), 671–682 (2005).
- [82] *I. Tavernelli, U.F. Röhrig, and U. Rothlisberger*, Molecular Physics **103**(6-8), 963–981 (2005).
- [83] *A. Dreuw and M. Head-Gordon*, Chemical reviews **105**(11), 4009–4037 (2005).
- [84] *B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez*, Molecular Physics **104**(5-7), 1039–1051 (2006).
- [85] *F. Cordova, L. J. Doriol, A. Ipatov, M. E. Casida, C. Filippi, and A. Vela*, Journal of Chemical Physics **127**(16) (2007).
- [86] *E. Teller*, Journal of Physical Chemistry **41**(1), 109–116 (1937).

- [87] G. Groenhof, L. V. Schäfer, M. Boggio-Pasqua, H. Grubmüller, and M. A. Robb, *Journal of the American Chemical Society* **130**(11), 3250–3251 (2008).
- [88] C. Burmesiter, PhD thesis, Physics, University of Göttingen (2013).
- [89] G. E. O. Borgstahl, D. R. Williams, and E. D. Getzoff, *Biochemistry* **34**(19), 6278–6287 (1995).
- [90] M. Kumauchi, M. T. Hara, P. Stalcup, A. Xie, and W. D. Hoff, *Photochemistry and photobiology* **84**(4), 956–969 (2008).
- [91] S. Devanathan, S. Lin, M. A. Cusanovich, N. Woodbury, and G. Tollin, *Biophysical journal* **81**(4), 2314–2319 (2001).
- [92] A. Losi, T. Gensch, M. A. Van Der Horst, K. J. Hellingwerf, and S. E. Braslavsky, *Physical Chemistry Chemical Physics* **7**(10), 2229–2236 (2005).
- [93] C. P. Joshi, H. Otto, D. Hoersch, T. E. Meyer, M. A. Cusanovich, and M. P. Heyn, *Biochemistry* **48**(42), 9980–9993 (2009).
- [94] K. Saito and H. Ishikita, *Proceedings of the National Academy of Sciences of the United States of America* **109**(1), 167–172 (2012).
- [95] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman, *Journal of Computational Chemistry* **24**(16), 1999–2012 (2003).
- [96] M. J. Harms, C. A. Castañeda, J. L. Schlessman, G. R. Sue, D. G. Isom, B. R. Cannon, and B. García-Moreno E., *Journal of Molecular Biology* **389**(1), 34–47 (2009).
- [97] M. J. Harms, J. L. Schlessman, G. R. Sue, and E. Bertrand García-Moreno, *Proceedings of the National Academy of Sciences of the United States of America* **108**(47), 18954–18959 (2011).
- [98] Isborn. C. M., A. W. Götz, M. A. Clark, R. C. Walker, and T. J. Martínez, *Journal of Chemical Theory and Computation* (2012).
- [99] A. Laio, J. VandeVondele, and U. Rothlisberger, *Journal of Chemical Physics* **116**(16), 6941–6947 (2002).

- [100] D. Das, K. P. Eurenium, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodoscek, and B. R. Brooks, *Journal of Chemical Physics* **117**(23), 10534–10547 (2002).
- [101] K. Senthilkumar, J. I. Mujika, K. E. Ranaghan, F. R. Manby, A. J. Mulholland, and J. N. Harvey, *Journal of the Royal Society Interface* **5**(SUPPL. 3), S207–S216 (2008).
- [102] D. J. Sindhikara, N. Yoshida, M. Kataoka, and F. Hirata, *Journal of Molecular Liquids* **164**(1-2), 120–122 (2011).
- [103] W S Lazarus-Barlow, **20**, 145–157 (1896).
- [104] K. J. Hellingwerf, J. Hendriks, and T. Gensch, *Journal of Physical Chemistry A* **107**(8), 1082–1094 (2003).
- [105] Z. Salamon, T.E. Meyer, and G. Tollin, *Biophysical Journal* **68**(2 I), 648–654 (1995).
- [106] J. Hendriks, T. Gensch, L. Hviid, M. A. Van Der Horst, K. J. Hellingwerf, and J. J. Van Thor, *Biophysical journal* **82**(3), 1632–1643 (2002).
- [107] J. E. Del Bene, W. B. Person, and K. Szczepaniak, *Journal of Physical Chemistry* **99**(27), 10705–10707 (1995).
- [108] M. Lozynski, D. Rusinska-Roszak, and H. . Mack, *Journal of Physical Chemistry A* **102**(17), 2899–2903 (1998).
- [109] H. G. Bohr, K. J. Jalkanen, M. Elstner, K. Frimand, and S. Suhai, *Chemical Physics* **246**(1-3), 13–36 (1999).
- [110] W. . Han, M. Elstner, K. J. Jalkanen, T. Frauenheim, and S. Suhai, *International Journal of Quantum Chemistry* **78**(6), 459–479 (2000).
- [111] O. V. Shishkin, L. Gorb, A. V. Luzanov, M. Elstner, S. Suhai, and J. Leszczynski, *Journal of Molecular Structure: THEOCHEM* **625**, 295–303 (2003).
- [112] R. Nakamura, N. Hamada, K. Abe, and M. Yoshizawa, *Journal of Physical Chemistry B* **116**(51), 14768–14775 (2012).
- [113] E. V. Gromov, I. Burghardt, H. Köppel, and L. S. Cederbaum, *Journal of Photochemistry and Photobiology A: Chemistry* **234**, 123–134 (2012).

- [114] *F. Schotte, H. S. Cho, V. R. I. Kaila, H. Kamikubo, N. Dashdorj, E. R. Henry, T. J. Graber, R. Henning, M. Wulff, G. Hummer, M. Kataoka, and P. A. Anfinrud*, Proceedings of the National Academy of Sciences of the United States of America **109**(47), 19256–19261 (2012).
- [115] *Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Srajer, and H. Ihee*, Nature chemistry **5**(3), 212–220 (2013).
- [116] *L. J. G. W. Van Wilderen, M. A. Van Der Horst, I. H. M. Van Stokkum, K. J. Hellingwerf, R. Van Grondelle, and M. L. Groot*, Proceedings of the National Academy of Sciences of the United States of America **103**(41), 15050–15055 (2006).
- [117] *R. Brudler, R. Rammelsberg, T. T. Woo, E. D. Getzoff, and K. Gerwert*, Nature structural biology **8**(3), 265–270 (2001).
- [118] *M. Unno, M. Kumauchi, J. Sasaki, F. Tokunaga, and S. Yamauchi*, Biochemistry **41**(17), 5668–5674 (2002).
- [119] *S. Devanathan, A. Pacheco, L. Ujj, M. Cusanovich, G. Tollin, S. Lin, and N. Woodbury*, Biophysical journal **77**(2), 1017–1023 (1999).
- [120] *N. Mataga, H. Chosrowjan, Y. Shibata, Y. Imamoto, and F. Tokunaga*, Journal of Physical Chemistry B **104**(21), 5191–5199 (2000).
- [121] *A. Espagne, P. Changenet-Barret, J. . Baudin, P. Plaza, and M. M. Martin*, Journal of Photochemistry and Photobiology A: Chemistry **185**(2-3), 245–252 (2007).
- [122] *P. Changenet-Barret, F. Lacombat, and P. Plaza*, Journal of Photochemistry and Photobiology A: Chemistry **234**, 171–180 (2012).
- [123] *M. Boggio-Pasqua, C. F. Burmeister, M. A. Robb, and G. Groenhof*, Physical Chemistry Chemical Physics **14**(22), 7912–7928 (2012).
- [124] *H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma*, Journal of Physical Chemistry **91**(24), 6269–6271 (1987).
- [125] *D. Morozov*, unpublished data, Computational Biomolecular Chemistry, Nanoscience Center, University of Jyväskylä.
- [126] *M.S. Gordon, M.A. Freitag, P. Bandyopadhyay, J.H. Jensen, V. Kairys, and W.J. Stevens*, Journal of Physical Chemistry A **105**(2), 306–307 (2001).